Tampere University

Alma Yrjänäinen

BIOCHEMICAL AND BIOPHYSICAL CHARACTERIZATION OF SALIVA- AND MILK HUMAN CARBONIC ANHYDRASE VI

Maidon ja syljen hiilihappoanhydraasi VI:n biokemiallinen ja biofysikaalinen karakterisointi

Faculty of Medicine and Life Sciences
Master's thesis
May 2019

ABSTRACT

Alma Yrjänäinen: BIOCHEMICAL AND BIOPHYSICAL CHARACTERIZATION OF SALIVA- AND MILK HUMAN CARBONIC ANHYDRASE VI
Master's thesis
Tampere University
Master's degree programme in Biomedical Technology, specialization in Cell Technology
June 2019

Background and aims of study: Carbonic anhydrase VI, CA VI, belongs to carbonic anhydrase isoenzymes that is known to secrete from serous acinar cells in the salivary glands and mammary gland to saliva and milk, respectively. The molecular weight of CA VI has been assessed with SDS-PAGE-electrophoresis as well as with Western blotting, yet not with liquid chromatography and light scattering. Also, the differences in the amino acid chains of milk and salivary CA VI are not thoroughly assessed. The aim of this study was to determine the molecular weight and particle size of milk and salivary CA VI in the physiological buffer. In addition, potential glycosylation differences and the melting temperature of CA VI were desired to study.

Methods: The molecular weight and particle size of CA VI isoenzymes were analyzed with SDS-PAGE, size-exclusion chromatography and light scattering, both static and dynamic, respectively. The melting temperature for salivary CA VI was determined with differential scanning calorimetry. The mass spectrometric analyses were not completed due to the malfunction of the apparatus.

Results: Based on SLS and standardized molecular weight determination, milk CA VI was shown to have the molecular weight of 203,815 kD and salivary CA VI was measured to be 204,509 kD. The hydrodynamic diameter for salivary CA VI was determined to be 20,3 ± 1,3 nm and the melting point was shown to be in the range of 50-55°C.

Conclusions: It can be concluded that both milk and salivary CA VI exist as oligomeric proteins in the physiological buffer. The pentameric state is probably the most prevalent form of assembly, which has not been reported earlier for any human CA isoenzyme. The existence of such a unique form opens new avenues for elucidating the prevalence and functional role of oligomerization in CA VI.

Keywords: carbonic anhydrases, carbonic anhydrase VI, size-exclusion chromatography, static light scattering, dynamic light scattering

Master's Thesis

Place	Faculty of medicine and Life Sciences, University of Tampere
Author	YRJANAINEN, ALMA MEDEA
Title	Biochemical and biophysical characterization of saliva- and milk human carbonic anhydrase VI
	Maidon ja syljen hiilihappoanhydraasi VI:n biokemiallinen ja biofysikaalinen karakterisointi
Pages	52
Supervisor	Professor Seppo Parkkila
Reviewers	Associate professor Vesa Hytönen and professor Seppo Parkkila
Date	March 25th 2019

Tiivistelmä

Taustat ja tavoitteet: Hiilihappoanhydraasi VI, CA VI, on hiilihappoanhydraaseihin kuuluva erittyvä isoentsyymi, jonka tiedetään ihmisellä erittyvän sylkirauhasen serooseista rauhassoluista sylkeen sekä rinnan maitorauhasen soluista äidinmaitoon. CA VI:n kokoa on tutkittu SDS-PAGE-elektroforeesilla sekä Western blottauksella, mutta ei nestekromatografian ja valonsironnan keinoin. Aiemmin ei ole myöskään tutkittu mahdollisia eroja syljen ja maidon CA VI:n aminohappoketjujen sokeriosissa, eli glykosylaatioissa. Tämän tutkimuksen tavoitteena oli määrittää maidon ja syljen hiilihappoanhydraasi VI:n molekyylipaino ja koko fysiologisessa puskurissa. Lisäksi haluttiin selvittää maidon ja syljen hiilihappoanhydraasi VI:n aminohappoketjujen mahdollisia sokeriosien eroja sekä tutkia CA VI:n sulamislämpöä.

Metodit: CA VI-isoentsyymien molekyylipainoa ja kokoa tutkittiin niin SDS-PAGE-elektroforeesilla kuin nestekromatografialla, sekä staattisella ja dynaamisella valonsironnalla. Sulamislämpö syljen CA VI:lle määritettiin differentiaaliskannauskalorimetrialla. Massaspektrometrisia analyyseja ei pystytty täydellisesti suorittamaan laitevian vuoksi.

Tulokset: Staattisen valonsironnan ja standardoidun molekyylipainomäärityksen perusteella maidon CA VI:n molekyylimassa oli 203,815 kD ja syljen CA VI:n oli 204,509 kD. Syljen CA VI:n halkaisijan mitattiin olevan 20,3 ± 1,3 nm ja sulamislämmön olevan 50-55°C asteen välillä.

Johtopäätökset: Voidaan todeta sekä maidon että syljen hiilihappoanhydraasi VI ovat fysiologisessa puskurissa oligomeerisiä, pentameerisen muodon ollessa todennäköisin. CA VI:n oligomeerista muotoa ei ole raportortoitu ennen. Sen olemassaolo mahdollistaa pentameerin mielekkäät jatkotutkimukset toiminnallisuuden selvittämiseksi.

Abstract

Background and aims of study: Carbonic anhydrase VI, CA VI, belongs to carbonic anhydrase isoenzymes that is known to secrete from serous acinar cells in the salivary glands and mammary gland to saliva and milk, respectively. The molecular weight of CA VI has been assessed with SDS-PAGE-electrophoresis as well as with Western blotting, yet not with liquid chromatography and light scattering. Also, the differences in the amino acid chains of milk and salivary CA VI are not thoroughly assessed. The aim of this study was to determine the molecular weight and particle size of milk and salivary CA VI in the physiological buffer. In addition, potential glycosylation differences and the melting temperature of CA VI were desired to study.

Methods: The molecular weight and particle size of CA VI isoenzymes were analyzed with SDS-PAGE, size-exclusion chromatography and light scattering, both static and dynamic, respectively. The melting temperature for salivary CA VI was determined with differential scanning calorimetry. The mass spectrometric analyses were not completed due to the malfunction of the apparatus.

Results: Based on SLS and standardized molecular weight determination, milk CA VI was shown to have the molecular weight of 203,815 kD and salivary CA VI was measured to be 204,509 kD. The hydrodynamic diameter for salivary CA VI was determined to be 20,3 ± 1,3 nm and the melting point was shown to be in the range of 50-55°C.

Conclusions:

It can be concluded that both milk and salivary CA VI exist as oligomeric proteins in the physiological buffer. The pentameric state is probably the most prevalent form of assembly, which has not been reported earlier for any human CA isoenzyme. The existence of such a unique form opens new avenues for elucidating the prevalence and functional role of oligomerization in CA VI.

Acknowledgements

The preceding year of planning my first, individual study plan, performing my experiments and finally writing my Master's thesis has been an unforgettable journey. For the past few months I have learned a great deal of scientific research that has encouraged me to continue this field.

Individually, I would like to thank Aulikki Lehmus, Niklas Kähkönen and Marianne Kuuslahti for their never-ending patience for guiding me in the essential experiments of my laboratory experiments. Also, I would like to thank Latifeh Azizi for extensive help provided during my HPLC analyses as well as DLS experiments. In addition, I would like to thank Harlan Barker for offering important assistance during my work. I would also like to thank Associate Professor Vesa Hytönen for his guidance as well as reviewing my Master's Thesis. Lastly, I would like to thank my supervisor Professor Seppo Parkkila for introducing me to carbonic anhydrases during my annual summer job as a research assistant in Tissue biology research group. His ceaseless passion for studying carbonic anhydrases has taught me resilience and humbleness towards scientific research and inspired me to continue despite the challenges.

Finally, thanks to those friends and family who have supported me during my work. My dear friend Michèle for daily support during challenging moments. Lastly, I would like to thank my wife Johanna for loving me as a person as I am, unconditionally and without a doubt.

Table of contents

Abbreviations

CA	Carbonic anhydrase
S_N2	Reaction mechanism of nucleophilic substitution
CARP	Carbonic anhydrase-related protein
gustin	Alternative name for carbonic anhydrase VI
cAMP	Cyclic adenosine monophosphate
PDEase	Phosphodiesterase enzyme
DMFT-index	Decayed, missing or filled teeth index
PROP	Propylithiouracil, study item
pSS	Primary Sjögren's syndrome
PTX	Pentraxin
GalNAc-4-SO_4	Galactose-N-acetylgalactosamine-4-SO_4 receptor
RT-PCR	Reverse transcriptase PCR
ELISA	Enzyme-linked immunosorbent assay
q-PCR	Quantitative PCR
HPLC	High-performance liquid chromatography
SEC	Size exclusion chromatography
SLS	Static light scattering
RALS	Right angle light scattering
DLS	Dynamic light scattering
PDI	Polydispersity index
MS	Mass spectrometry

Introduction

Carbonic anhydrases, CAs, are fundamental enzymes catalyzing the hydration of carbon dioxide (Chegwidden and Carter, 2000). There are 16 isoforms of α-CAs that are characterized in humans (Kalinin *et al.*, 2016). Among them is CA VI, the only secreted isoenzyme, that is produced by serous acinar cells in parotid and submandibular glands and human mammary glands resulting in secretion into saliva and human milk, respectively (Parkkila *et al.*, 1990; Karhumaa *et al.*, 2001). CA VI is known to possess anti-caries functionality as it balances the pH in the dental plaque microenvironment (Kimoto *et al.*, 2006). It has also been shown to have single nucleotide polymorphism resulting in alteration in bitter taste perception (Padiglia *et al.*, 2010; Feeney and Hayes, 2014). CA VI is also used as a biomarker in primary Sjögren's syndrome (Pertovaara *et al.*, 2011). It has also been stated that CA would have nerve growth factor-like characteristics as it might function as a trophic factor in the infant alimentary tract (Henkin *et al.*, 1999a; Karhumaa *et al.*, 2001).

Although milk and salivary CA VI isoenzymes have been studied parallelly with SDS-PAGE and Western blotting, there are no studies focusing solely on comparing both of the CA VI isoenzymes. As the amino acid sequences were analyzed with amino acid databases, they revealed 100% identity with human salivary CA VI, yet the sequenced polypeptides covered only 40% of the full-length CA VI sequence (Karhumaa *et al.*, 2001). Thus, there still remains a possibility that the milk and salivary CA VI isoenzymes have structural differences. Considering the different pattern of polypeptide bands for milk and salivary CA VI in the SDS-PAGE gel, there might be variation in protein glycosylation as well. To date, the particle size of CA VI has not been reported. Prior data has suggested that human CA VI might possess oligomeric assemblies (Pilka *et al.*, 2012), it was tempting to hypothesize that those structures exist, and they could be observed with the chosen methods in the current study.

This study was conducted to provide additional information for molecular weight and particle size of milk and salivary CA VI as well as assessing the potential post-translational differences between milk and salivary CA VI.

Literature review

Introduction to carbonic anhydrases

Carbonic anhydrases, CAs, are essential enzymes catalyzing the hydration of carbon dioxide $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$(Chegwidden and Carter, 2000). The reaction mechanism can be categorized as S_N2, nucleophilic substitution, when the Zn^{2+}-bound hydroxide group attacks on CO_2 followed by the bicarbonate displacement with water and finally, deprotonation of water to regenerate the hydroxide (Chegwidden and Carter, 2000). In all isoforms, the active site is located in the funnel-shaped cavity extending from the center of the protein to the surface (Stams and Christianson, 2000). The catalytic zinc atom is located at the bottom of this cavity coordinated by three histidine residues (Stams and Christianson, 2000). First, histidine ligands structurally stabilize the zinc-binding site and secondly, functionally maintain the reactivity of zinc-bound solvent for catalysis (Stams and Christianson, 2000).

The superfamily of carbonic anhydrases

To date, the superfamily of carbonic anhydrases consists of seven, genetically unrelated families of isoforms noted as α-, β-, γ-, δ-, ζ-, η- and θ-CAs (Vullo *et al.*, 2017). The distribution of the carbonic anhydrase families is presented in figure 1. α-CAs are present solely in vertebrates (Chegwidden and Carter, 2000). This α-family of isoforms is also present in many algae, plants and eubacteria (Chegwidden and Carter, 2000). In addition to α-CAs, plants and algae also encode for β-, γ-, δ- and θ-CAs (Vullo *et al.*, 2017). Marine diatoms, forming a major group of algae, encode for ζ-CAs (Innocenti *et al.*, 2010). CAs identified in bacteria belong to the α-, β- and γ-CA-classes (Vullo *et al.*, 2017).

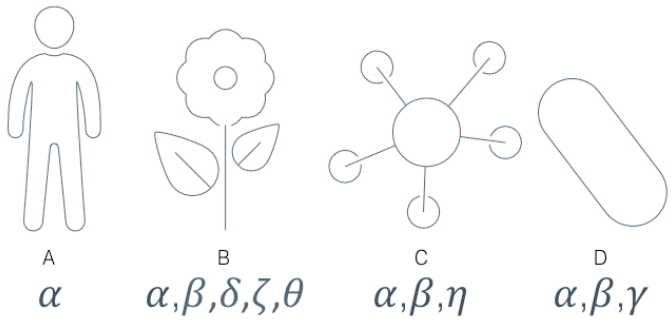


Figure 1 Schematic distribution of CA families in A) animals, B) plants and algae, C) protozoa and D) bacteria.

The remaining three isoforms are classified as carbonic anhydrase related proteins, CARPs, that are catalytically inactive, because these isoforms lack histidine residues forming the coordination bond to zinc atom (Aspatwar *et al.*, 2010). Nevertheless, CARPs have a high sequence and structure similarity with other isoforms and are thus considered as a part of CA isoform family (Tashian *et al.*, 2000). To date, three CARPs have been discovered: CARP VIII, CARP X and CARP XI (Tashian *et al.*, 2000).

Distribution and functions of α-carbonic anhydrases

There are 16 isoforms of α-CAs, which have been characterized in humans (Kalinin *et al.*, 2016). Histidine residues, His-94, His-96 and His-119, are conserved within mammalian α-CA isoforms, which stabilize the hydration of CO_2 and catalysis of water and, thus, are important for hydration of CO_2 (Stams and Christianson, 2000). These isoforms are involved in a variety of diverse physiological processes, such as respiration, acid-base balance, bone resorption and calcification, as this conversion from CO_2 to bicarbonate is essential in animals (Chegwidden and Carter, 2000). Also, α-CA-isoenzymes take part in many other biological pathways including ion, gas and fluid transfer (Chegwidden and Carter, 2000). The distribution of the enzymatically active isoforms is presented in a schematic mammalian cell model (Fig. 2). There are five cytosolic CA-isoforms (CA I, CA II, CA III, CA VII and CA XIII), five membrane-bound isoforms (CA IV, CA IX, CA XII, CA XIV and CA XV), two mitochondrial isoforms (CA VA and CA VB) and finally, the secretory isoform CA VI that is found in saliva and milk (Innocenti, Scozzafava and Supuran, 2010).

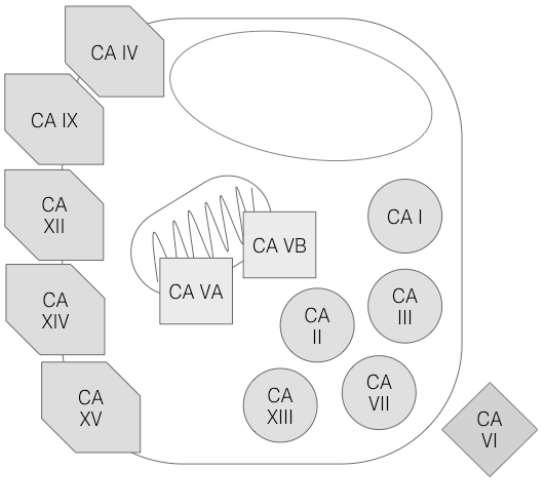


Figure 2 Distribution of α-CAs in an animal cell model. Cytosolic isoforms are CA I, CA II, CA III, CA VII and CA XIII. Membrane-bound isoforms are CA IV, CA IX, CA XII, CA XIV and CA XV. CA VA and CA VB are mitochondrial isoforms. To date, CA VI is found to be the only secreted isoform.

The subcellular location and functions for mammalian α-CAs are described subsequently, except secretory CA VI, which is introduced in the chapter 'Carbonic anhydrase VI'.

Cytosolic carbonic anhydrases

Cytosolic CAs, CA I and CA II, are found in erythrocytes as they function as classical catalysts in carbon dioxide hydration (Frost, 2014). Physiologically, CA II has a higher activity compared with CA I, and thus can be considered more relevant in actual CO_2-hydration process (Frost, 2014). CA II is also widely present in kidneys (Brown *et al.*, 1983). When investigating immunohistochemically stained rat tissue, CA II was observed in intercalated cells and in proximal tubules, the loop of Henle and principal cells found in collecting ducts (Brown *et al.*, 1983). In addition, the uneven distribution of CA II due to the heterogenic development rate of the nephrons will partly explain the reduced capacity for urinary acidification, which is reported for infants born prematurely (Lönnerholm and Wistrand, 1983). CA III, however, differs from the other isoenzymes protecting cells from irreversible protein oxidation, as it has shown to have protective role against oxidative stress (Thomas and Mallis, 2001). CA III is present in skeletal muscle tissue and both of adipose tissues (Frost, 2014). Human CA VII was identified through genomic screening (Montgomery *et al.*, 1991) and was later discovered to have a long, predominant form and a shorter form (Bootorabi *et al.*, 2010). It was also localized in the colon, liver and skeletal muscle, and to some extent in the brain (Bootorabi *et al.*, 2010). CA XIII was identified in 2004 and localized to be residing in various human tissues, including thymus, small intestine, spleen, prostate, ovary, colon and testis (Lehtonen *et al.*, 2004).

Mitochondrial carbonic anhydrases

CA VA and CA VB are localized in the mitochondria with varying expression in different tissues (Shah *et al.*, 2000). CA VA was found to express only in liver, skeletal muscle and kidney, CA VB, conversely, was detected in most tissues (Shah *et al.*, 2000). When inhibiting CA VA and VB, it has been found to reduce lipogenesis through a novel mechanism and these isoenzymes could thus be considered having potential for therapeutic applications (Poulsen *et al.*, 2008).

Membrane-associated carbonic anhydrases

The membrane associated CAs in human include CA IV, CA IX, CA XII, CA XIV and finally CA XV (Innocenti, Scozzafava and Supuran, 2010). CA IV is shown to be present in the lung, kidney, heart, brain, the capillary of the eye and erythrocytes (Frost, 2014). It was also claimed to form a functional complex with Na^+/bicarbonate co-transporter NBC1 to

maintain correct pH balance in retinal environment (Yang *et al.*, 2005). Furthermore, recent evidence suggests that CA IV would also have a role in skin wound healing, as CA IV significantly accelerated wound re-epithelialization when introduced to the wound in zebra fish (Barker *et al.*, 2017). CA IX and XII are limited in normal expression but are found to be tumor-related isoforms of α-CAs (Frost, 2014). For example, CA IX interacts with pH-regulating sodium bicarbonate co-transporter (NBC1) and anion exchanger 2 (AE2) that are components of migration apparatus (Svastova *et al.*, 2012). Thus, CA IX was shown to actively contribute to cell migration that is one of the factors promoting the metastatic cascade (Svastova *et al.*, 2012). CA XII expression, however, has been showed to be in correlation with the positive prognosis in breast cancer (Barnett *et al.*, 2008). When studying mRNA of CA XIV, it was highly expressed in brain and weaker signal was seen in colon, small intestine, urinary bladder, and kidney (Frost, 2014). The newest member of the membrane associated CAs, CA XV, was shown to be present in mouse kidney (Tolvanen *et al.*, 2013). It provided an attractive example of adaptation in the terms of evolution, as the functions of CA XV seems to be taken over by CA IV (Tolvanen *et al.*, 2013).

Secreted carbonic anhydrases

To date, only one secretory carbonic anhydrase isoenzyme, CA VI, has been distinguished. As this thesis is primary concentrating on CA VI, I will describe the structure and functions of human CA VI and non-human CA VI in more detail in the following chapter.

Carbonic anhydrase VI

Carbonic anhydrase VI is a secreted enzyme that was first found in the late 1940's (Rapp, 1946). To date, CA VI is the only secretory isoenzyme found in mammals (Patrikainen *et al.*, 2016). As the literature of CA VI shows, the secretory isoenzyme has also been called *'gustin'* until it was proven to be the same protein as CA VI in the late 1990's (Thatcher *et al.*, 1998). Nevertheless, the name *'gustin'* is still used occasionally when referring to CA VI in the scientific literature (Padiglia *et al.*, 2010; Calò *et al.*, 2011; Melis *et al.*, 2013; Barbarossa *et al.*, 2015). Although CA VI has been carefully studied in the past 60 years, the fundamental functions of the secretory isoenzyme still remain uncovered.

In the following chapter, I will introduce the general properties considering human CA VI; its origin and locations in the human body and secretional alterations. Then, I will discuss the catalytic properties of CA VI as functional enzyme. Also, I will briefly introduce the key biological phenomena that CA VI is related in; dental caries, polymorphism, taste perception and finally Sjögren's disease. I will also describe the gene and protein structure of CA VI.

In the second chapter I will introduce the essential findings of non-human CA VI; where CA VI is located in various animals and what are the main functions reported on CA VI; the dental caries model in mice, bitter taste perception studies by using mice and finally, the kidney disease model in pig. Then, I will briefly cover the theoretical aspects of used methods in my thesis.

human carbonic anhydrase VI

Human carbonic anhydrase VI is produced by serous acinar cells in parotid and submandibular glands resulting in secretion into saliva (Parkkila *et al.*, 1990). It was later demonstrated that serous salivary glands, von Ebner's glands, located in the posterior tongue among lingual muscle fibers, deliver CA VI to the immediate proximity of taste buds (Leinonen *et al.*, 2001). This local secretion implies the connection between CA VI and taste bud function (Leinonen *et al.*, 2001). The relation between CA VI and taste perception is described more detail in chapter '' Carbonic anhydrase VI in taste perception''.

CA VI was first purified from human saliva in 1987 and considered a distinct form of a novel isoenzyme although CA VI was found to be genetically and immunologically related to CA II (Murakami and Sly, 1987). The secretion of CA VI into saliva is found to follow circadian periodicity together with α-amylase, even though it was independent of salivary secretion (Parkkila *et al.*, 1995). When CA VI was measured from the human serum, the concentration varied greatly within individuals and circadian rhythm was not as evident compared with salivary CA VI (Kivelä *et al.*, 1997). As CA VI is abundant in human parotid saliva, CA VI isoenzyme was first thought to have an effect on regulating pH in saliva (Parkkila *et al.*, 1990), but was later found to have no significant correlation with pH buffering capacity, suggesting that CA VI is not directly regulating salivary pH (Kivelä *et al.*, 1999).

In human serum, CA VI was demonstrated to bound into IgG (Kivelä *et al.*, 1997). This could prevent CA VI from protein degradation or alternatively, function as a targeting factor to guide CA VI to cells not expressing this isoenzyme (Kivelä *et al.*, 1997). CA VI has been studied within different populations of both sexes. CA VI concentration in saliva was found to be somewhat lower in young women than young men (Kivelä *et al.*, 1997; Kivelä *et al.*, 2003). CA VI production has also been studied in pregnant women (Kivelä *et al.*, 2003). Pregnancy did not affect CA VI concentration in saliva even though the buffer capacity of saliva was discovered to be lower than in non-pregnant women (Kivelä *et al.*, 2003). This further confirms that CA VI is not directly regulating salivary pH (Leinonen *et al.*, 1999). CA VI was also found to secrete from the mammary glands to human milk (Karhumaa *et al.*, 2001). When measuring the concentration of CA VI in human colostrum, the first secretion of mammary glands after giving birth, it was discovered to have a CA VI-concentration

approximately eight times higher compared with the concentration measured in mature milk collected 90 days postpartum (Karhumaa *et al.*, 2001). The high concentration of colostral CA VI could recompense the low salivary CA VI in the infant alimentary tract and their incomplete salivary secretion (Karhumaa *et al.*, 2001). CA VI is found to have similar characteristics than nerve growth factor NGF, thus this isoenzyme might function as a trophic growth factor in the infant alimentary tract (Karhumaa *et al.*, 2001).

The genetic structure of carbonic anhydrase VI

CA 6 gene is located at chromosome 1 on the tip of the short arm (Sutherland *et al.*, 1989). The sequence analysis has been executed for *CA 6* gene in order to investigate its evolution (Patrikainen *et al.*, 2017). Phylogenetic analysis showed that CA IX, XII, XIV and CA VI share a common ancestor proposing that the given isoenzymes appeared as a result of two whole genome duplications in the evolution of early vertebrates (Patrikainen *et al.*, 2017). As CA VI evolution was more carefully considered, a likely evolutive pathway was introduced resulting in varying CA VI domain structures (Patrikainen *et al.*, 2017). First, it is supposed that exon coding for the cytoplasmic domain in ancestral CA VI was replaced by PTX-encoding exon (Patrikainen *et al.*, 2017). Simultaneously, the transmembrane helix present in the common ancestor CA was transformed into an amphipathic helix changing the conformation of the evolving CA VI (Patrikainen *et al.*, 2017). Later, as the PTX domain was lost, CA VI reached its final structure having an amphipathic helix at the C-terminus (Patrikainen *et al.*, 2017). These conclusions were supported by four essential observations; exon lengths of transmembrane-helix-coding exon, the loss of the cytoplasmic domain in ancestor CA VI and PTX domain in mammalian lineage, rearrangement in glucose transporters near PTX domain and, finally, the abundance of PTX domain in non-mammalian CA VI (Patrikainen *et al.*, 2017). This sequence of structural changes has ultimately resulted in the unique structure of CA VI consisting of catalytic CA domain and a amphipathic helix lacking the cytoplasmic domain (Patrikainen *et al.*, 2017). However, this structural analysis has been performed by studying CA VI found in zebra fish that ultimately possesses a pentameric assembly with the pentraxin domain.

The amino acid structure of carbonic anhydrase VI

Human CA VI has a signal peptide that is first cleaved resulting in the mature form of CA VI; a 291-amino acids containing protein showing strong sequence similarity particularly to membrane bound isoenzymes (Supuran and De Simone, 2015). There are distinct properties in both amino- and carboxy-terminals. The amino-terminal region possesses a conserved active site, N-glycosylated site and cysteine residues forming a conserved disulphide bond (Supuran and De Simone, 2015). The carboxy-terminal region contains

both a unique sequence of 30 amino acids only found in CA VI as well as a recognition determinant for glycosyltransferases (Supuran and De Simone, 2015).

Human CA VI has a molecular weight of 42 kD determined from saliva (Murakami and Sly, 1987). Later, CA VI isoenzymes, both salivary and milk, have been determined as 42-kD proteins by Western blotting (Karhumaa *et al.*, 2001). However, it was reported that deglycosylation of both salivary and milk CA VI, reduced their molecular size to 36 kD (Karhumaa *et al.*, 2001). It was reported having two N-linked oligosaccharide side chains, an oligosaccharide linked into the amide nitrogen of asparagine residue (Leinonen *et al.*, 2001). These side chains terminate with GalNAc-4-SO_4 (Hooper *et al.,* 1995). The nonglycosylated milk CA VI isoenzyme was digested and sequenced with matrix-assisted laser desorption/mass spectrometry and ultimately, shown to have 100% identity compared with salivary CA VI and the coverage being 40% of the full-length CA VI (Karhumaa *et al.*, 2001). It was speculated that as the milk CA VI amino acid sequence was only 40% covered of salivary CA VI, there is a possibility that salivary CA VI and milk CA VI might possess alternating amino acid structures (Karhumaa *et al.*, 2001).

The crystallized structure of the catalytic domain of CA VI, aa 32-279, has been published at 1.9 Å resolution (Pilka *et al.*, 2012). The catalytic domain, including the amino acids 21-290 (Homo sapiens carbonic anhydrase VI mRNA, GenBank entry 21706434), was subcloned into a vector with His-tag and after transforming, culturing, lyzing and purifying the protein, His-tag was removed from the catalytic domain (Pilka *et al.*, 2012). When crystallized, it was shown to assemble as orthorhombic space group $P2_12_12_1$, P describing the lattice being a primitive, rectangular shape and $2_12_12_1$ depicting the symmetry group of its configuration in three-dimensional space (Pilka *et al.*, 2012).

The catalytic domain of human CA VI was found to have two molecules in a single asymmetric unit (Pilka *et al.*, 2012). The asymmetric unit forms a dimer structure, as the active sites of the domains are juxtaposed towards the dimeric center, facing each other (Pilka *et al.*, 2012). 11 hydrogen bonds are formed at the dimeric interface, burying a statistically significant area of accessible surface that exceeds the average for biological dimers (Pilka *et al.*, 2012). When analyzed with size-exclusion chromatography, two equally-sized peaks were detected and thus argued to be the dimeric and monomeric assemblies of the catalytic domain of the recombinant CA VI (Pilka *et al.*, 2012). This novel information about the dimeric occurrence of CA VI could explain its functionality *in vivo*, like it has been hypothesized for CA IX (Alterio *et al.*, 2009) and CA XII (Whittington *et al.*, 2004). However, it has to be emphasized that only the catalytic domain, aa 32-279, were crystallized (Pilka *et al.*, 2012). For CA IX, the dimerization is mediated by Cys 41-Cys 41-disulphide bond that interconnects two monomers (Alterio *et al.*, 2009). Due to the

orientation of the monomers, both active sites exposed to extracellular space for efficient CO_2-hydration (Alterio *et al.*, 2009). This is consistent with the tissues reported having expression or over-expression of the CA IX (Alterio *et al.*, 2009). For CA XII, the catalytic domain is also found to obtain a dimeric structure (Whittington *et al.*, 2004). Dimerization is predicted to occur in the transmembrane helices as they contain the dimerization motifs GXXXG and GXXXS (Whittington *et al.*, 2004). Still, only the catalytic, globular domains of CA XII and CA IX have been able to crystallize leaving the whole structure, thus the external structures undiscovered (Whittington *et al.*, 2004; Alterio *et al.*, 2009).

The dimer formation has been previously reported on sheep, when ovine tear secretion was analyzed with Western analysis and CA VI was assumed to be the in dimeric state considering its fitting molecular weight (Ogawa *et al.*, 2002). However, other methods were not used to verify this assumption. Also, an alternative form of CA VI, a CCAAT/Enhancer-Binding Protein Homologous Protein-dependent stress-inducible form (CHOP-inducible form), has been reported on murine cell culture studies (Sok *et al.*, 1999; Matthews *et al.*, 2014). In response to stress, an alternative CA VI, CA VI-b is encoded and retained in the intracellular space (Sok *et al.*, 1999). CA VI-b was sequenced and found to lack a signal peptide and, thus, is not predicted to encode a secretory protein (Sok *et al.*, 1999). Still, as the CA VI-b was sequenced to contain all the known residues contributing to the active site and most of the residues were conserved among other CAs, it can be assumed that it is very likely for CA VI-b to have carbonic anhydrase and esterase activities (Sok *et al.*, 1999). Activation of endoplasmic reticulum unfolded protein stress response pathway leads to expression of the CA VI-b from unknown promoter (Matthews *et al.*, 2014). This unknown promoter is regulated by CHOP as it acts as a downstream regulator activating proapoptotic effectors (Sok *et al.*, 1999). Thus, a CHOP-regulated induction of CA VI-b was thought to be proapoptotic (Sok *et al.*, 1999). Later, however, it was reported that CA VI-b solely has little influence on cell viability (Matthews *et al.*, 2014). CA VI-b was reported functioning as a part of pro-survival branch of the CHOP-signaling but still, it is suggested that the role of CA VI-b is more complex in cell death decisions (Matthews *et al.*, 2014). For more specific functions for CA VI-b, it was shown that CA VI-b is necessary for mediating the beneficial properties of the brain derived neurotrophic factor BDNF in hypoxic neuron survival (Matthews *et al.*, 2014).

The domain was found to assemble in the canonical CA-fold having several, short α-helices and ß-sheets surrounding the ten-stranded ß-sheet located in the center (Pilka *et al.*, 2012). It has a conserved, intramolecular disulphide-linkage between cysteines 42 and 224 that stabilizes the active site, so that the Zn^{2+}-bound hydroxide is oriented for catalytic reaction (Pilka *et al.*, 2012). The disulphide linkage identical to CA VI can be found also in the other CA-structures (Pilka *et al.*, 2012). When CA VI was aligned with CA I-CA XIV, it was

predicted to share high sequence identity with membrane-bound isoenzymes CA IV and CA XIV in addition to cytosolic isoenzyme CA II (Pilka *et al.*, 2012). Due to sequence insertions and deletions, the backbone of CA VI was found to differ slightly from other CA-isoenzymes as three varying regions were identified (Pilka *et al.*, 2012).

Catalytic properties of CA VI

For human CA VI, both reactions, hydration CO_2 and dehydration of bicarbonate, have been investigated in terms of their specific activity rates. Human CA VI has a maximum value of $2.9 \times 10^7 \frac{1}{Ms}$ for kinetic constant $\frac{k_{cat}}{K_M}$ when dehydrating bicarbonate ions are conveyed into CO_2 and water (Leinonen, 2008). The half-saturated concentration of bicarbonate, K_M, is significantly lower when compared with other CAs having the K_M<20 mM (Leinonen, 2008). However, it resembles the concentration of bicarbonate in saliva and milk that possibly suggests the capacity for CA VI to catalyze bicarbonate with high k_{cat}, respectively (Leinonen, 2008).

When considering the hydration reaction, the catalytic efficiency, $\frac{k_{cat}}{K_M}$, for human CA VI is $4.9 \times 10^7 \frac{1}{Ms}$, which is one-third of the fast, cytosolic CA II, but K_M of has a value close to other CAs, with few exceptions (Nishimori *et al.*, 2007). Its turnover rate k_{cat} is $3.4 \times 10^5 \frac{1}{s}$, which is a somewhat moderate value in comparison with other CAs (Nishimori *et al.*, 2007). Uniquely assembled amino acid residues explain the activity of CA VI; threonine and serine residues near the conserved proton shuttle residue His64 might be bulkier compared with alanine, present in highly efficient CA II in the given location, so that Ser/Thr residues might result in interfering in the movement of His64 (Nishimori *et al.*, 2007). CA VI has residues typical for other CAs as well, which are involved in binding activators or inhibitors (Nishimori *et al.*, 2007). Also, CA VI has characteristic residues near zinc ligands affecting its activity (Nishimori *et al.*, 2007). Interestingly, CA VI is found to be the most sensitive to bicarbonate and chloride inhibition suggesting towards the biomolecular environment of the given enzyme containing alterable amounts of proteins and anions (Nishimori *et al.*, 2007).

Human CA VI is found to activate cAMP PDEase at the physiological concentration, but the mechanism differs from Ca^{2+}-dependent-calmodulin activation (Law *et al.*, 1987). The PDEase activation by CA VI might explain how CA VI affects taste function, as cAMP PDEase is an important regulator in taste perception (Law *et al.*, 1987). However, this research has not been further continued and additional data has not been discovered.

Recently, the melting temperature of salivary CA VI has also been studied (Kazokaite *et al.*, 2015). For native, salivary CA VI the melting point was reported being +53,6 °C by inhibitor binding method Fluorescence Thermal Shift Assay, FTSA, which determines thermal stability via fluorescent probes while the temperature is steadily increased (Kazokaite *et al.*, 2015).

Carbonic anhydrase VI and caries

Human CA VI has been found to have anti-caries characteristics; caries-free children expressed higher CA activity than the children having active caries (Szabó, 1974). Also, negative correlation between CA VI concentration and poor oral hygiene was discovered by investigating the number of decayed, missing or filled teeth (DMFT index) within individuals (Kivelä *et al.*, 1999). Low CA VI concentration was measured with individuals having a low DMFT index value (Kivelä *et al.*, 1999). Salivary CA VI was found to be a natural component of enamel pellicle and shown to be functional and active on the enamel surface (Leinonen *et al.*, 1999). CA VI is suggested converting bacterial plaque-derived H^+-ions with salivary bicarbonate into carbon dioxide and water, and thus preventing the acidification of the microenvironment (Leinonen *et al.*, 1999). As the phenomena were further investigated, salivary CA VI was inhibited with acetazolamide resulting in the lower pH value of dental plaque (Kimoto *et al.*, 2006). This further suggests that salivary CA VI penetrates the plaque and neutralizes acidic H^+-ions with salivary bicarbonate in the given microenvironment and thus, can be considered as an anti-caries protein (Kimoto *et al.*, 2006).

Polymorphism of carbonic anhydrase VI

As the nuanced functions of CA VI still remain unclear, studies have also been carried out to discover whether single nucleotide polymorphisms (SNPs) occur in the coding region of *CA 6* (Peres *et al.*, 2010; Aidar *et al.*, 2013). Three SNPs have been reported being present at the coding region of *CA 6* (Peres *et al.*, 2010). Buffer capacity or caries susceptibility had no statistically significant difference between the groups (Peres *et al.*, 2010). However, a positive association between the buffer capacity and rs2274327 (C/T) polymorphism was observed when allele T and genotype TT were significantly less frequent among individuals with high buffer capacity (Peres *et al.*, 2010).

It was suggested that instead of coding for C allele resulting in threonine, allele T results in coding Met that ultimately disrupts a highly conserved short ß-sheet (Peres *et al.*, 2010). This change is likely to interfere with the function of CA VI, like buffer capacity (Peres *et al.*, 2010). Genotype TT of rs2274327 (C/T) polymorphism was also associated with

significantly lower salivary CA VI concentration (Aidar *et al.*, 2013). Still, the enzymatic activity of CA VI showed no significant differences between groups with altering polymorphisms (Aidar *et al.*, 2013). It is proposed that as rs2274327 allele T lies at the site of O-glycosylation, it is very likely to have this site glycosylated (Aidar *et al.*, 2013). O-glycosylation has been shown to be a protective signal against proteasomal degradation (Guinez *et al.*, 2008).

Carbonic anhydrase VI in taste perception

Carbonic anhydrase VI has been considered as a salivary growth factor that maintains the function of taste buds (Henkin *et al.*, 1999a). Inhibition of CA VI synthesis is correlated with the abnormal development of taste buds resulting in the loss of taste function (Henkin *et al.*, 1999a). By treating with exogenous zinc, CA VI content was increased in parotid saliva and in addition, patients exhibited improvement in taste and smell acuity (Henkin *et al.*, 1999b). It was discussed that zinc would act as a gene transcription stimulant resulting in the synthesis and secretion of CA VI (Henkin *et al.*, 1999b). It is suggested that CA VI acts as a trophic factor promoting growth and development of taste buds while affecting taste bud stem cells (Henkin *et al.*, 1999b). The long-term exposure of low pH of epithelial cells has been found to lead to apoptosis, so it is suggested that CA VI may protect taste buds' taste receptor cells (TRCs) from apoptosis by regulating their pH (Leinonen *et al.*, 2001). Proton-gated potassium channels are inactivated because of acid stimuli depolarizing the TCRs (Kinnamon and Roper, 1987). CA VI might neutralize excess H^+-ions in the taste bud microenvironment thus accelerating the recovery of proton-gated potassium channels (Leinonen *et al.*, 2001).

SNPs of *CA 6* have been investigated in terms of altering taste perception (Padiglia *et al.*, 2010). SNP in *CA 6* coding region, rs2274333 (A/G) resulting in Ser90Gly substitution in amino acid sequence, has been investigated whether it would have an effect on bitter taste perception (Padiglia *et al.*, 2010). The SNP rs 2274333 affecting exon 3 of the *CA 6* gene seems to be altering the capacity of CA VI to bind zinc (Padiglia *et al.*, 2010). The genotype AA, associated with fully functional protein, was statistically shown to be associated with high responsiveness to PROP, a well-known substance for tasting the bitter, and controversially, individuals expressing GG, associated with a disruptive form of CA VI, would have the lowest responsiveness for PROP (Padiglia *et al.*, 2010). Additionally, the alteration within CA VI may be related to differences in papillae densities and oral chemosensory abilities among alternating phenotypic groups of PROP (Padiglia *et al.*, 2010). However, no significant differences were observed in fungiform papillae density considering any SNPs of *CA 6* (Feeney and Hayes, 2014).

The connection between PROP and CA VI has been debated. Another study suggested that CA 6 polymorphism would associate with salt perception rather than PROP (Feeney and Hayes, 2014). It has been shown that saltiness intensity increases as a function of PROP intensity, thus indicating that the ratio between PROP and salt might not be the optimal choice to define PROP functionality (Feeney and Hayes, 2014).

Carbonic anhydrase VI as a marker for Sjögren's syndrome

Primary Sjögren's syndrome (pSS) is a chronic rheumatic disease having the characteristic symptoms of dryness of the eyes and mouth, and possible extraglandular symptoms (Pertovaara *et al.*, 2011). In pSS, abundant autoantibody production occurs as a sign of autoimmunity and it can be clinically determined by observing renal manifestations (Pertovaara *et al.*, 2011). As anti-CA antibodies have been suggested being pathogenic when involved in the formation of renal manifestations, anti-CA antibodies were measured together with serum creatinine, sodium and potassium as well as urinary excretion substances (Pertovaara *et al.*, 2011). The inverse correlation between the concentration of anti-CA-antibodies in the serum and serum sodium levels could imply that CA VI is one of the significant regulators of acid-base homeostasis in pSS patients (Pertovaara *et al.*, 2011). Moreover, anti-CA-antibodies, including anti-CA VI, correlated significantly with urinary pH in pSS patients, thus suggesting that these anti-CA-antibodies might be associated in renal acidification capacity (Pertovaara *et al.*, 2011). Detection of autoantibodies to CA VI in the early stage of pSS might imply the diagnostic relevance that CA VI could provide, in addition to the markers Ro and La (Matossian and Micucci, 2016).

Localization of non-human carbonic anhydrase VI

Various species have been used for localizing CA VI within different tissues. To date, CA VI has been purified from the saliva and parotid glands of several mammals, like sheep (Fernley *et al.*, 1979), rats (Feldstein and Silverman, 1984), humans (Murakami and Sly, 1987), cattle (Asari *et al.*, 2000), swine (Nishita *et al.*, 2001), mice (Kimoto *et al.*, 2004) and dogs (Kasuya *et al.*, 2007).

In sheep, measurable amounts of CA VI were found in parotid, submandibular and salivary glands by radioimmunoassay (Fernley *et al.*, 1988). In rat, CA VI was discovered immunohistochemically to be present in the serous acinar cells, ductal cells and ductal content of von Ebner's glands and in the demilune, ductal cells plus ductal content of rat lingual mucous glands (Leinonen *et al.*, 2001). Moreover, CA VI was found to be present in the serous acinar cells and duct cells of tracheobronchial glands and in the secretory cells and the base of the ciliated cells of the tracheobronchial surface epithelium in addition to

Clara cells residing on the bronchiolar surface epithelium (Leinonen *et al.*, 2004). When considering the presence of CA VI in cattle, it was shown by RT-PCR and immunohistochemical methods that bovine mammary gland together with salivary and lacrimal glands synthetize CA VI and, thus it is also found in cow milk and tear fluid (Kitade *et al.*, 2003). Also, positive signs were seen in the bovine stratified epithelium of esophagus, in the stratified epithelium of bovine forestomach, the mucous cells of upper glandular region of the large intestine, and finally, in some ductal cells of submandibular, monostomatic sublingual and esophageal glands (Kaseda *et al.*, 2006). Afterwards, it was shown by RT-PCR and ELISA that CA VI is present also in bovine liver (Nishita *et al.*, 2007). In swine, however, CA VI was determined by ELISA to be present in saliva, parotid gland, sublingual gland and submaxillary gland, but in serum, bile, seminal plasma, and finally, gall bladder, too, which was observed immunohistochemically to be present at the columnar epithelial cells lining the gallbladder (Nishita *et al.*, 2011). Still, CA VI measured from serum and gallbladder was only a small fraction compared with the concentration measured in parotid gland demonstrating that the highest and most significant concentration was found in parotid gland (~440 μg/1 g wet tissue), and moderate concentrations (~100 μg/1 g wet tissue) pancreas, lung, bile, seminal plasma and colostrum (Nishita *et al.*, 2011). Also, CA VI was localized in the epithelial cells of the renal straight distal tubules (Nishita, Yatsu, Murakami, *et al.*, 2014). In mice, CA VI was immunohistochemically observed in acinar cells of, in the duct contents of the anterior gland of nasal septum, in the lateral nasal gland and ultimately, in the mucus covering the respiratory and olfactory mucosa and in the lumen of the nasolacrimal duct (Kimoto *et al.*, 2004). By using immunohistochemical methods and both RT-PCR and q-PCR, canine CA VI was localized in mucosal epithelial cells, in the cytoplasm of serous acinar and ductal epithelial cells of the nasal mucosa and glands including the vestibule of the nose (Kasuya *et al.*, 2007).

In addition to the mammalian animal models, CA VI has been characterized in pufferfish *Tagifugu rubripes*, and by comparing semi-quantitative RT-PCR and q-PCR results simultaneously, the highest expression for CA VI was found in liver, and after that kidney, heart, blood and gill also showed moderate expression levels (Sumi *et al.*, 2018). By investigating liver tissue more carefully by using fluorescence-in-situ-hybridization FISH, the expression might be seen within hepatocytes (Sumi *et al.*, 2018). However, it has been shown that in the case of zebra fish, the non-mammalian CA VI is present with a pentraxin (PTX) domain, an inflammation regulating multimeric protein complex (Patrikainen *et al.*, 2017). CA VI binding PTX is most closely related to SAP and CRP, that are fundamental fluid-phase pattern recognition molecules functioning in innate immune system reactions (Bottazzi *et al.*, 2016). When the distribution of CAVI-PTX in zebrafish was studied, immunohistochemical staining showed a high presence of CA VI in skin, heart, gills, and

the swim bladder (Patrikainen *et al.*, 2017). In detail, CA VI was observed on the cell surface (Patrikainen *et al.*, 2017).

Functions of carbonic anhydrase VI in non-human species

To date, published non-human CA VI studies are somewhat focused on measuring and localizing CA VI (Fernley *et al.*, 1979; Feldstein and Silverman, 1984; Fernley *et al.*, 1988; Asari *et al.*, 2000; Karhumaa *et al.*, 2001; Leinonen *et al.*, 2001; Nishita *et al.*, 2001, 2011; Kasuya *et al.*, 2007; Patrikainen *et al.*, 2017; Sumi *et al.*, 2018). However, several studies show a variety of physiological phenomena that CA VI is part of. Transcriptomic responses and histological alterations in the submandibular gland, the stomach and duodenum of *CA 6*-deficient mice have been reported (Pan *et al.*, 2011). This knock-out mouse model for CA VI deficiency has been useful to demonstrate changes in gene expression morphology in lower airways (Patrikainen *et al.*, 2016) and similarly, *CA 6* -deficient mice were used on studying caries development induced by *Streptococcus mutans* (Culp *et al.*, 2011). A knock-down zebra fish model was established for determining behavioral and morphological changes due to transient lack of CA VI (Patrikainen *et al.*, 2017).

Still, there are certain physiological phenomena associated with CA VI that have been investigated in the animal model. *CA 6*-deficient mice were used to examine bitter taste perception (Patrikainen *et al.*, 2014). Although no morphological changes were detected in the routine light microscopy of fungiform papillae, *CA 6*-deficient mice had altered behavior for bitter taste (Patrikainen *et al.*, 2014). Compared with the wild type, behavioral monitoring showed that *CA 6*-deficient mice preferred bitter quinine solution (Patrikainen *et al.*, 2014). This suggests that CA VI might have a role in bitter taste avoidance (Patrikainen *et al.*, 2014). They also discovered that Car6-/- mice had a lower consumption of salty NaCl-solution compared with the wild type (Patrikainen *et al.*, 2014).

Another mice model study showed how CA VI was also found to be dependent on Zn^{2+}T4-mediated Zn^{2+}-transport, a direct Zn^{2+} transport from the cytoplasm into the *trans*-Golgi network, to maintain its stability (McCormick and Kelleher, 2012).

As a secretory protein, CA VI is transferred via exocytosis to the mucus on olfactory and respiratory mucosa from acinar cells of nasal mucosa (Kimoto *et al.*, 2004). CA VI secretion was detected to be dependent on vesicle associated membrane protein 8, VAMP8 (Wang *et al.*, 2007). By using a VAMP8-knock-out line of mice, they were able to show immunohistochemically and by EM analysis that acinar cells had evenly distributed CA VI within the cytoplasm compared with the wild type that had CA VI restricted to the apical

side of the acinar cells (Wang *et al.*, 2007). This suggests that VAMP8 has a key role in regulating secreted proteins in murine acinar cells (Wang *et al.*, 2007).

CA VI has also been considered as a biomarker for swine kidney diseases (Nishita *et al.*, 2014). By measuring the levels of CA VI from diseased and wild type pigs by ELISA and executing a statistical analysis of Dunnett's multiple comparison test, it was postulated that CA VI could serve as a potential biomarker for kidney diseases as it was increased when the epithelial cells of the renal straight distal tubules were damaged (Nishita *et al.*, 2014).

Used methods for determining biochemical and biophysical characteristics of CA VI

In the next chapter, I will briefly explain the theory of the study methods used in this thesis and why these methods were chosen to study human CA VI. First, I will discuss the affinity chromatography for protein purification. Then, I will continue with theoretical introduction to high performance liquid chromatography as well as to static and direct light scattering.

Affinity chromatography for protein purification

The specific affinity chromatography used for CA VI purification was developed originally for CA II purification (Wolpert *et al.*, 1977) and subsequently it has been used in numerous CA VI studies (Law *et al.*, 1987; Parkkila *et al.*, 1990; Parkkila *et al.*, 1995; Karhumaa *et al.*, 2001; Leinonen *et al.*, 2001). Since it is considered as commonly accepted practice among researchers, this method was chosen to purify native CA VI from saliva and milk.

High performance liquid chromatography

High performance liquid chromatography (HPLC) is a separation method for a mixture of different compounds where the separation method is chosen due to the nature of the studied sample. Considering the aim of the study, to compare the structures of human salivary and milk CA VI, size-exclusion was selected as a suitable method for investigating the structural properties of the enzymes obtained from both sources. Size-exclusion chromatography is based on the porous structure of the stationary phase beads, which will adsorb the smaller molecules thus elongating their retention time (Moldoveanu and David, 2013). The larger molecules are not retained as they cannot penetrate into the pores and are, thus, eluted earlier (Moldoveanu and David, 2013). Based on the relation of protein's molecular size, the Stokes radius and the shape of the particle, elution volume is increased

as the particle size decreases (Slotboom *et al.*, 2008). SEC is commonly used for determining the molecular weight of a studied particle (Slotboom *et al.*, 2008).

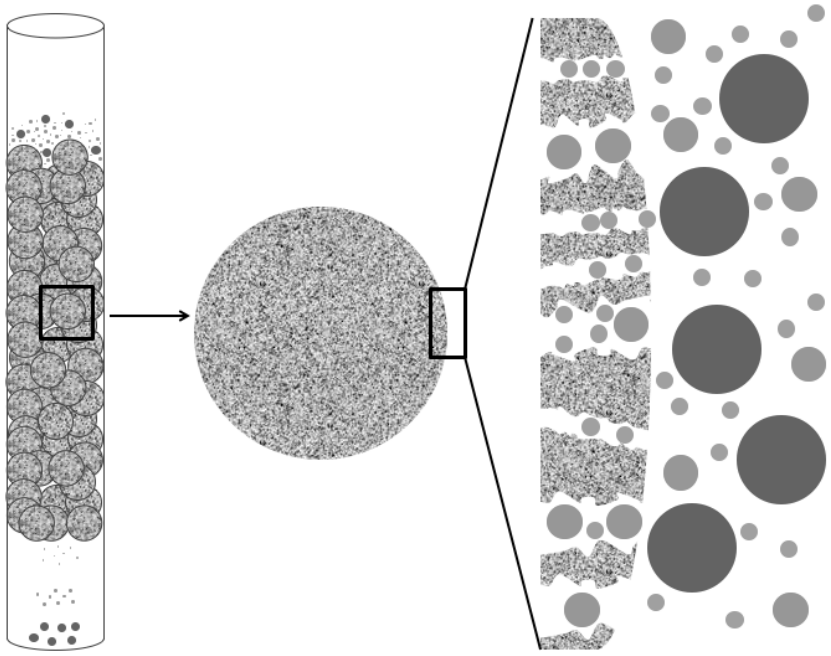


Figure 3 The schematic description of the size-exclusion process. The column packed with chemically and physically inert beads, which are porous and thus adsorbing smaller particles of the sample. This stationary phase separates sample molecules according to their size, as the largest particles have the shortest retention time and the smallest particles having the longest retention time.

In SEC analysis, the flow rate of the sample, the pressure of the column and temperature can be adjusted to meet the required physical and chemical circumstances. Separation of the sample particles can be improved as the flow rate is decreased (Netopilík, 2017). The separation power of the studied particles will increase since molecules are able to compensate for the decreased rate of transversal diffusion (Netopilík, 2017). Also, the equilibrium state of the molecules in and outside of the porous beads of the stationary phase is more properly reached since they are given enough time to set (Netopilík, 2017). To overcome the resistance offered by the packed beads in the column, a suitable pressure must be provided to the mobile phase (Striegel *et al.*, 2009). Generally, constant-flow reciprocating pumps are used as the constant-pressure pumps can suffer from changes affecting the column backpressure, like temperature-derived flow variations within the column (Striegel *et al.*, 2009). Other specifications central to the pumping system are repeatability, short-term precision, pump pulsation, drift and flow rate accuracy (Striegel *et al.*, 2009). Repeatability depicts the ability of resetting the pump to the exact flow rate repetitively, short term precision depicts how the precise and consistent is the volume output of the pump over time, pump pulsation occurs as operational functions like piston movement cause flow changes, drift measures the continuous changes in the pump output

over longer periods of time and finally, pumping accuracy reports the preciseness of the pump to deliver the exact flow rate that is set at the system (Striegel *et al.*, 2009).

Commonly in SEC, the temperature is increased to improve sample solubility (Striegel *et al.*, 2009). Still, many analysis are performed out of room temperature because of their nature; for example biological samples might be susceptible to degradation or disassembly at high temperatures (Striegel *et al.*, 2009). Within a good solvent, however, the particle size is changed very little with the temperature (Striegel *et al.*, 2009). Even though the temperature has only a minor influence on the slope and position of the molar mass calibration curve, these small curve shifts will ultimately have a significant effect on the accuracy of the results (Striegel *et al.*, 2009). Thus, large temperature fluctuations should be avoided in SEC analysis (Striegel *et al.*, 2009).

In the given study, LC-10Ai HPLC isocratic pump was used to maintain pressure during the analysis as it offers low pulsation performance and highly reproducible analyses (Shimadzu corporation, 2018).

Static light scattering

Static light scattering (SLS) measures the molecular weight of a macromolecule using the relationship between molecule's scattered light intensity and molecular weight and size (Malvern Instruments Ltd., 2014). SLS is based on the fundamental Rayleigh relationship,

$$\Delta LS = \left(\frac{I_\theta}{I_0}\right)_{protein} - \left(\frac{I_\theta}{I_0}\right)_{buffer} = K\left(\frac{dn}{dc}\right)^2 M_W C$$

in which $\frac{I_\theta}{I_0}$ is the ratio of scattered light at angle θ and the incident light for the sample, $\left(\frac{I_\theta}{I_0}\right)_{protein}$, and the buffer, $\left(\frac{I_\theta}{I_0}\right)_{buffer}$, are subtracted. K is a constant depending on the refractive index of a solution without studied particles present, the used wavelength and the scatter angle θ, and the distance between the studied particle and the detector; $\frac{dn}{dc}$ is a specific refractive index increment for the studied protein; M_W describing the molecular mass of the studied protein; C is the concentration of the protein (Slotboom *et al.*, 2008). These will be resulting in ΔLS, a parameter describing the excess scattered light due to the size of the studied protein (Slotboom *et al.*, 2008).

Molecular weight M_W of a studied macromolecule can be estimated in different ways (Slotboom *et al.*, 2008). In the two-detector method, the usage of parameter $\frac{dA}{dc}$ that relates changes in the UV absorbance intensity of the studied solution to the studied protein concentration, enables computing the molecular weight for the studied macromolecule (Slotboom *et al.*, 2008). Alternatively, the molecular weight of a studied macromolecule can also be retrieved by using retention volumes for standard proteins with known molecular weights (Slotboom *et al.*, 2008). After converting the molecular weights to a logarithmic scale they can be plotted by their specific retention volumes and, thus, unknown molecular mass can be calculated by its known retention volume (Malvern Instruments Ltd., 2014).

Right-angle light scattering (RALS) is an instrument measuring the intensity of scattered light at 90° to the incident light beam (Malvern Instruments Ltd., 2014). The molecular weight of the studied protein is directly calculated from measured intensity and sample concentration (Malvern Instruments Ltd., 2014). Due to 90° angle, any noise created by the change of the refractive index is minimized and, thus, it provides an excellent sensitivity and signal-to-noise ratio (Malvern Instruments Ltd., 2014). In addition, it is well-suited for estimating the size of proteins, generally having the diameter of <15 nm, as the measurement assumes that intensity is the same in both 0° and 90° angle (Malvern Instruments Ltd., 2014). For larger molecules, the assumption is incorrect as they display significant angular dependence considering the light that it scatters (Malvern Instruments Ltd., 2014).

Dynamic light scattering

Dynamic light scattering, DLS, is a method based on the essential property of light scattering, as the electromagnetic waves will induce electronic distortion that will be emitted in varying directions (Panchal *et al.*, 2014). The scattering of light is depending on multiple parameters; (1) the homogeneity of particles, (2) their refractive index relative to surrounding medium and (3) particle shape and size (Jonasz and Fournier, 2007). To estimate the sample homogeneity, the polydispersity index, PDI, compares the absolute standard deviation of the particle size distribution to the mean, creating the relative polydispersity (Nobbmann, 2014). The overall polydispersity is converted into polydispersity index PDI which is calculated as a square of light scattering polydispersity (Nobbmann, 2014). This is a value for estimating the degree of uniformity of a distribution; for an ideally uniform sample, polydispersity index PDI would be 0.0 (Nobbmann, 2014). The used apparatus in this study, Malvern Zetasizer Nano (Malvern Instruments Ltd., Worcestershire, UK), suggests that samples having a greater value than 0,7 are highly polydisperse and, thus, are perhaps unsuitable for DLS measurements (Malvern

Instruments Ltd., 2008). The range of 0,08-0,7 is considered the best operational area for the used algorithms in Malvern Zetasizer Nano apparatus (Shaw, 2018).

In DLS, particles' Brownian motion is measured and subsequently used for calculating the hydrodynamic diameter with Stokes-Einstein equation (Malvern Instruments Ltd., 2008). DLS illuminates particles with a laser so that intensity fluctuations of scattered light can be analyzed (Malvern Instruments Ltd., 2008). This type of scattering is referred to as quasielastic light scattering (Striegel *et al.*, 2009). In DLS, the time-dependent fluctuations of scattered light are measured, whereas in SLS time-averaged fluctuations of scattered light are of interest (Striegel *et al.*, 2009). In this thesis, the used DLS apparatus, Malvern's Zetasizer NZ uses a digital correlator to compare the speckle patterns of signal intensity of particular part at one point in time (Malvern Instruments Ltd., 2008). With similar Brownian motion the intensity correlation is high within time points (Malvern Instruments Ltd., 2008). As the Brownian motion of the larger particles are minor compared with smaller particles it can be detected as a slow fluctuation of the intensity of the speckle pattern compared to smaller and thus faster fluctuating particles (Malvern Instruments Ltd., 2008). Ultimately, the correlation function for different-sized particles can be measured and used for the size determination of the studied particle (Malvern Instruments Ltd., 2008). The diffusion coefficient and the hydrodynamic radius can be measured and, thus, the diameter of the studied particle can be computed (Panchal *et al.*, 2014). DLS analysis can be combined with SEC, in-line DLS, as it measures the particle size in the size-determined order, or it can be done solely using a separate DLS apparatus, batch DLS. Both types of DLS were used in this thesis.

Results of batch DLS can be presented as intensity, volume or number distributions (Malvern Instruments Ltd., 2008). The peak area of the intensity distribution describes the scattering intensity of the given particle (Malvern Instruments Ltd., 2008). Since the large particles are able to scatter more light than smaller particles due to greater surface area, the scattering intensity of a particle is proportional to the sixth power of the particle diameter (Malvern Instruments Ltd., 2008). Based on the intensity distribution, a volume distribution can be generated that can be further computed into number distribution (Malvern Instruments Ltd., 2008). Number distribution depicts the peak are by the particle prevalence of the sample, e.g. equally-sized peaks indicate to an equal number of both samples (Malvern Instruments Ltd., 2008). Volume distribution depicts peak area by the particle's volume (Malvern Instruments Ltd., 2008). Similar to intensity distribution, larger particles have greater volume than smaller particles, thus creating a larger peak area for greater particles and minor peak area for smaller particles (Malvern Instruments Ltd., 2008). In this study, volume distributions are analyzed for determining the fundamental size of the CA VI isoenzyme.

The hydrodynamic diameter will not depend solely on the core size of the particle (Shaw, 2018). It is also affected by the ionic concentration of the sample medium or particle's surface structures (Shaw, 2018). Low ionic concentration creates an electric double layer of ions that decreases diffusion speed and thus, leads to larger, apparent hydrodynamic diameter (Shaw, 2018). Additional surface structures will affect diffusion speed, too (Shaw, 2018). Adsorbed polymers projecting out from the particle surface will reduce the diffusion speed and also lead to apparently larger size (Shaw, 2018). Finally, for non-spherical particles, DLS will measure the diameter of a sphere having the same average translational diffusion coefficient as the non-spheroid particle (Shaw, 2018). Taken together, all of the previously described phenomena can change the apparent hydrodynamic diameter by several nanometers (Shaw, 2018).

For determining the melting point of a studied protein, an increase in hydrodynamic radius and scattering intensity indicates the presence of denatured protein aggregates (Malvern Instruments Ltd., 2008). Thus, batch DLS offers a sufficient method for measuring the stability of a studied protein since the temperature can be steadily increased to desired value (Malvern Instruments Ltd., 2008).

Objectives

The overall purpose of this research is to discover the potential oligomeric structure of milk and salivary CA VI. Also, it is desired to determine whether the human salivary and milk CA VI-proteins have differing glycosylations within their structures. In addition, the melting point for CA VI will be studied.

The used methods include:

- Affinity chromatography
- High-performance liquid chromatography with static and dynamic light scattering
- Batch dynamic light scattering
- Mass spectrometry

These methods should provide sufficient information for the assembly of human CA VI protein, and, to gain additional information of the essential structure of both salivary and milk CA VI, the melting point for the CA VI and, lastly, their potential differences considering post-translational modifications.

Materials and methods

Salivary CA VI

Salivary CA VI isoenzyme, for both HPLC-DLS analysis and batch DLS analysis, were collected in two batches due to insufficient amount of protein after the first batch was purified in terms of CA VI. All the used reagents, solutions prepared and used in the protocol are more carefully described in Appendix 1. Thus, only the given solution names are used in the 'Materials and methods' -section. For CA VI purification from milk and saliva, the original protocol of Murakami and Sly used in CA VI purification was followed (Murakami and W S Sly, 1987).

Saliva collecting

First, 100 ml and 300 ml of human saliva was collected for the purification of the CA VI isoenzyme. Volunteers expectorated into an ice-cold centrifuge tube, which contained 2 ml of collecting buffer (0,1 M TRIS, 0,2 M Na_2SO_4, 0,2 M Benzamidine, pH 8,70) for 35 ml of collection to reduce unspecific binding. After collection, until the total volume was collected, saliva was stored at +4 ℃.

For purification, 100 ml and 300 ml of saliva were centrifuged (16'000× g, 30 min, Sorvall Lynx 4000, Thermo Scientific, Madison, USA) for removing extraneous material. Then, saliva was diluted with Binding buffer (0,1 M TRIS and 0,2 M, pH 8,70) up to 0,7 liter and 1,5 liter, respectively. p-Aminomethylbenzenesulfonamide–Agarose (Sigma) for the affinity chromatography matrix were added 3 ml and 6 ml, respectively, and both were left on magnetic stirrer +4°C/ON.

Purification of salivary CA VI

The affinity chromatography matrix was collected with Milli-Pore filtration system (Millipore Corporation, Burlington, USA) with vacuum, followed by washing with 500 ml (100 ml×5 times) and 1000 ml (100 ml×10 times) of wash 1 buffer (0,1 M TRIS, 0,2 M Na_2SO_4, 0,2 M Benzamidine, 20% glycerol, pH 8,70), respectively. The following wash on each purification was carried out with 300 ml of the same buffer but at lower pH (7,00). The matrix was collected from the filtration system by pipetting into the chromatography column. Finally, CA VI was eluted in 2 ml fractions with the elution buffer (0,1 M TRIS, 0,4 M NaN_3, 1mM Benzamidine, 20% glycerol, pH 7,0). For maximizing the amount of CA VI in the first fraction, the first volume of the elution buffer was pipetted into the column, removed to an

external centrifuge tube and left on a 3D-stirrer for 30 minutes. After incubation, CA VI was eluted, and fractions of 2 ml were collected. Fractions were analyzed using SDS-PAGE-electrophoresis. The molecular weight of salivary CA VI was determined using Precision Plus Protein™ Standards Dual Color (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and molecular weight marker and salivary CA VI bands were visualized using the Colloidal Blue Staining Kit™ (Invitrogen). This purification protocol was repeated once, because after the first purification the amount of impurities observed in the SDS-PAGE gel was significant considering the further analysis of the protein. Concentration was measured using NanoDrop One (Thermo Scientific, Madison, USA).

Milk CA VI

500 ml and 1000 ml of human milk were donated by Tampere university hospital. After melting into liquid form, the milk samples were centrifuged (16'000× g, 30 min, Sorvall Lynx 4000, Thermo Scientific, Madison, USA). The solid, fatty layer was removed manually, and 300 ml and 750 ml of milk were diluted with the binding buffer to the total volume of 1 liter and 1,75 liter, respectively. 4 ml and 12 ml of p-Aminomethylbenzenesulfonamide–Agarose (Sigma) was added for each of the dilutions. For reducing the unspecific binding, 25 ml and 75 ml of the collecting buffer were added to the diluted milk samples, respectively. Both of the diluted samples were left on magnetic stirrer +4°C/ON.

Purification of milk CA VI

The same protocol was followed as in salivary CA VI purification. 1 liter and 3 liters of wash 1 buffer were used, following 250 ml and 1000 ml of wash 2 buffer used, respectively. The matrix was collected, and the first fraction was incubated and eluted as described earlier. All fractions were analyzed with SDS-PAGE electrophoresis. As well as for salivary CA VI, also the molecular weight of milk CA VI was determined using Precision Plus Protein™ Standards Dual Color (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and molecular weight marker and milk CA VI bands were visualized using the Colloidal Blue Staining Kit™ (Invitrogen). Like salivary CA VI, also milk CA VI was found to possess significant amount of impurities. Thus, milk CA VI fractions were also repurified to reduce the amount of impurities. Concentration was measured using NanoDrop One (Thermo Scientific, Madison, USA).

Sample preparation

Before further analysis, milk and salivary CA VI-samples were analyzed with SDS-PAGE to observe them in parallel manner. The molecular weight of milk and salivary CA VI was

determined using Precision Plus Protein™ Standards Dual Color (Bio-Rad Laboratories, Inc., Hercules, CA, USA) and molecular weight marker, milk and salivary CA VI bands were visualized using the Colloidal Blue Staining Kit™ (Invitrogen). After reaching a sufficient level of purity and measuring the concentration, the elution buffer in both milk CA VI and salivary CA VI were changed to suitable buffers, 150 mM NaCl and 50 mM TRIS for high-performance liquid chromatography, and 100mM ammonium acetate for mass spectrometry, respectively. The elution buffer was changed because of the relatively high concentration of NaN_3 with Vivaspin Turbo 15 (30 mm×118 mm, Sartorius AG, Göttingen, Germany) resulting in NaN_3-concentration being less than 10^{-9}M to minimize its light absorbing characteristics on the light scattering analysis. After changing the buffer, the concentration was measured using NanoDrop One (Thermo Scientific, Madison, USA).

High-performance liquid chromatography and dynamic light scattering experiments

The molecular weight of CA VI was determined with high-performance liquid chromatography (HPLC) Shimadzu system (CBM-20A, Shimadzu Corporation, Kyoto, Japan) with Malvern Zetasizer μV instrument (Malvern Instruments Ltd., Worcestershire, UK) by using static light scattering (SLS) and dynamic light scattering (DLS) methods. HPLC system was equipped with solvent delivery unit LC-10Ai, degassing unit DGU-20A5R, autosampler SIL-20A, and fluorescence detection RF-20Axs, UV/VIS detector SPD-20A and communications bus module CBM-20A and was used for analyzing both CA VI samples. The separation was carried out using a Superdex 200 5/150 GL column (10 mm×300 ml, particle size 13μm, GE Healthcare, Uppsala, Sweden) and was equilibrated with filtered 50 mM TRIS, 150 mM NaCl, pH 7,50 running buffer. The chromatographic runs were performed in the thermostatic cabin and detector and column temperature were set to +12°C with a flow rate of 0,1 $\frac{ml}{min}$. The sample volume was calculated so that 50 μg of both of the CA VI proteins and 30 μg of standard proteins were injected by using an autoinjector. In both runs, the detector and column temperature were Data were processed using Lab Solution version 5.51 (Shimadzu Co., Kyoto, Japan) and OmniSec 4.7 (Malvern Instruments Ltd., Worcestershire, UK) softwares. Molecular weight determination was performed two ways; (1) by calibrating the light scattering detector using the monomeric peak of bovine serum albumin and light scattering intensity and (2) by the size-exclusion chromatography standard curves of four standard proteins (Sigma-Aldrich, Inc., Saint Louis, MO, USA); bovine serum albumin 66,7 kD, alcohol dehydrogenase 150 kD, beta-amylase 200 kD and carbonic anhydrase II 29 kD.

Batch-dynamic light scattering

To verify the in-line DLS measurement, an individual measurement for determining the hydrodynamic diameter of milk and salivary CA VI, 0,9 $\frac{mg}{ml}$ and 0,97 $\frac{mg}{ml}$, respectively, was performed with batch-DLS-analysis using Malvern Zetasizer Nano (Malvern Instruments Ltd., Worcestershire, UK) equipped with 4mW laser lamp (λ=633 nm). 100 μl of both samples were analyzed, and five parallel measurements were calculated at the temperature of +25 °C. Data were processed using Zetasizer 7.03 software (Malvern Instruments Ltd., Worcestershire, UK). Average values of hydrodynamic diameter, d.nm, were computed by five individual measurements for both samples.

Furthermore, the hydrodynamic diameter of salivary CA VI was analyzed in increasing temperature. First, sample was centrifuged (13'000× g, 5 min) filtered with 0,2 μm filter and diluted into lower concentration of 0,28 $\frac{mg}{ml}$. Subsequent protocol was used: first, sample was equilibrated in each temperature for 10 minutes. Then, five parallel measurements were performed on each desired temperature; +20 °C, +25 °C, +30 °C, +37 °C, +40 °C, +45 °C, +50 °C, +55 °C, +60 °C, +65 °C and +70 °C. This experiment was conducted for assessing the changes in hydrodynamic diameter as temperature is altered and determine the temperature range for the melting point of CA VI.

Results

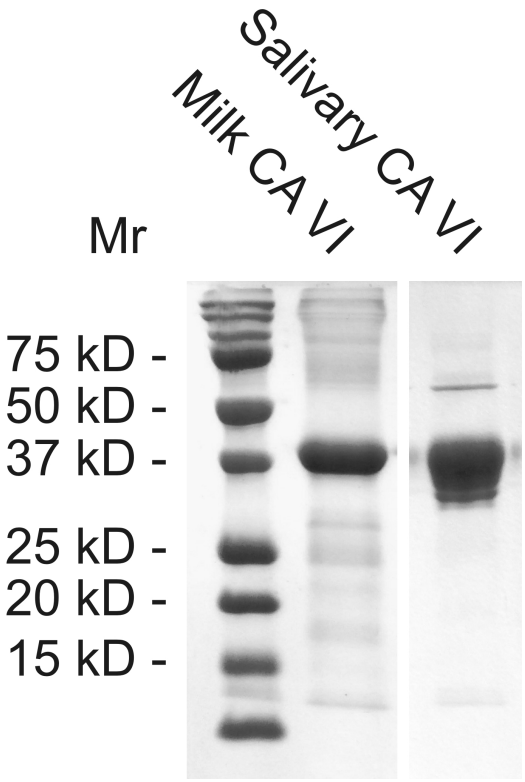
To date, the parameter defining the specific UV sensitivity as a concentration derivative, dA/dC, has not been reported in the literature. Thus, dA/dC for CA VI was computed using Protparam tool (Gasteiger *et al.*, 2005). Carbonic anhydrase VI (CAH6_HUMAN; protein ID: P23280) protein sequence was selectively used from position 18 to position 308 to model native CA VI properties. Based on the chosen sequence resembling the intact protein without the signal peptide sequence (1-17 aa), dA/dC was calculated and further used in the computations of HPLC software OmniSec 4.7 (Malvern Instruments Ltd., Worcestershire, UK). Additionally, the theoretical value of CA VI isoenzyme's molecular weight was retrieved from ProtParam tool to provide a theoretical reference value for SLS-measured molecular weights by OmniSec 4.7.

Table 1 Computed values for CA VI by ProtParam tool.

Computed data by ProtParam tool	
Theoretical MW	33,56952 kD
dA/dC	1,666 ml/g

Purification of CA VI

Both salivary and milk CA VI were purified twice due to significant impurities present after the first purification. Approximately 1l of human milk resulted in 1,630 $\frac{mg}{ml}$ of CA VI. For salivary CA VI, the total of 400 ml of saliva resulted in 2 $\frac{mg}{ml}$ of CA VI. Milk and salivary CAVI were analyzed in parallel with 10% SDS-PAGE gel. In picture 1 below, 11 µg of milk CA VI and 10 µg of salivary CA VI are presented. The purified milk CA VI showed a single, strong band close to the anticipated size, approximately 42 kD, in SDS-PAGE-gel among a small number of weaker bands both larger and smaller in size. Salivary CA VI showed a wider band around 42-36 kD.



Picture 1 The SDS-PAGE of native human milk and salivary CA VI. Left: molecular weight standards presented as kilodaltons (kD). Middle: purified native human milk CA VI Right: purified native salivary CA VI. For both CA VI-proteins, molecular mass calculated from mobility resulting in approximately 42 kD.

High-performance liquid chromatography and dynamic light scattering

The molecular size of milk and salivary CA VI were estimated after liquid chromatography by SLS and DLS, respectively. The results are presented individually in the following subchapters.

HPLC-DLS-analysis of milk CA VI

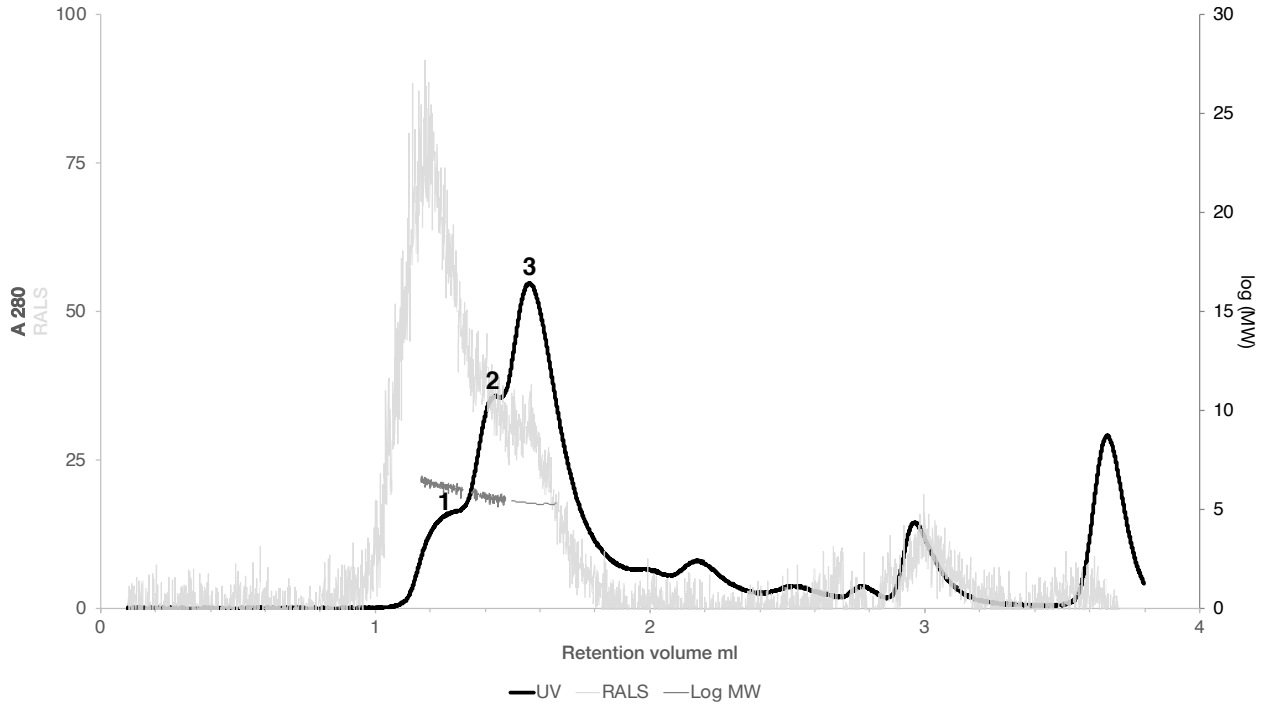
HPLC-analysis resulted in approximate estimation of molecular weight of the milk CA VI isoenzyme. Also, the hydrophobic radius of milk CA VI was measured using DLS, both in-line and batch, respectively. Data retrieved from the static light scattering (SLS), the standardized molecular weight and the direct light scattering (DLS) are presented in table 2, along with the provided theoretical value for the molecular weight of CA VI.

Table 2 HPLC-DLS results of milk human CA VI. In the first section, theoretical molecular weight for the monomeric human CA VI (hCA VI) protein is provided by ProtParam. In the middle section SLS and in-line DLS results, molecular weight calculated by SLS, standardized molecular weight and hydrodynamic radius, are shown for peak 1-3 in the given retention volumes. In the last section, batch DLS-measurement for hydrodynamic diameter is shown.

Sample	Theoretical MW (kD)		SEC ret. volume (ml)	SLS MW (kD)	Stand. MW (kD)	Rh DLS in-line (nm)	Batch DLS (d.nm)
milk hCA VI	33,56952	Peak 1	1,318	1642,000	414,686	N/A	15,25
		Peak 2	1,474	427,265	250,094	N/A	
		Peak 3	**1,562**	**203,815**	**188,025**	**217,21**	

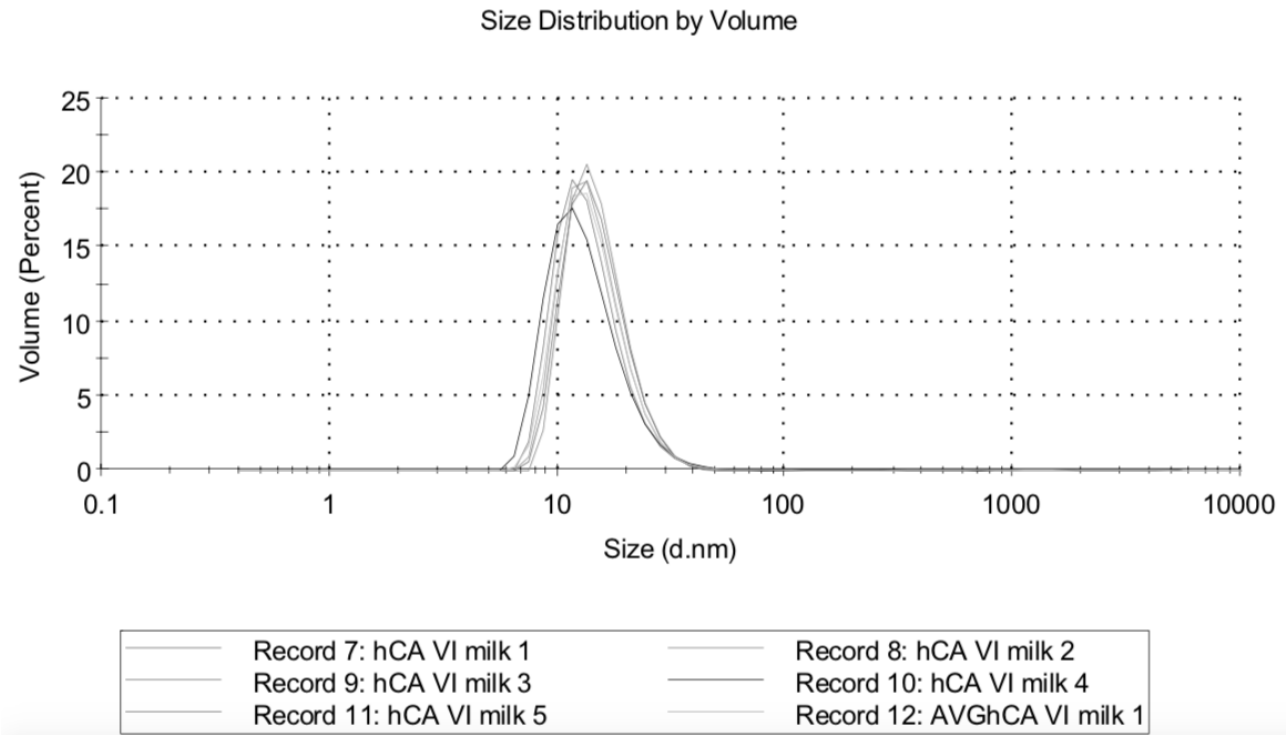
Graph 1 shows UV absorbance (A280), light scattering (RALS) and logarithmic value of molecular weight (log MW) as a function of retention volume (ml) for milk CA VI. The main peak of milk CA VI, marked as peak 3, was eluted in 1,56 ml retention volume according to UV absorbance A280, shown as a black curve in graph 1. However, SLS intensity peak differed from its shape when compared with UV absorbance in milk CA VI as it is depicted in light grey. SLS intensity for the main peak predicted a molecular weight of 203,815 kD for milk CA VI, which is depicted as a near-horizontal dark grey line across the peak 3. There are also other molecular weight estimates shown in the graph; at retention volume $V_{ret,1}$=1,318 ml, the molecular weight estimate is approximately 1642 kD, and $V_{ret,2}$=1,474 ml the molecular weight of the particles is estimated to be 427 kD.

In-line DLS data for the eluted peak was also collected, but due to its significantly high values and standard deviation, the average of the hydrodynamic radius being 217,21 nm and standard deviation 39,79 (data not shown), it is not shown in the graph.



Graph 1 Assessment of the oligomeric size of milk human CA VI. Size-exclusion chromatography was used to study the properties of human native, milk CA VI. The X-axis shows the retention volume for the analyzed sample. The left Y-axis shows the UV absorption intensity of 280 nm wavelength (A280, black curve) and Right Angle Light Scattering intensity (RALS, light grey). In the right Y-axis, logarithmic values for molecular weight were computed by light scattering intensity. In-line DLS-measurement for hydrodynamic radius Rh is not shown.

As the hydrodynamic radius was unsuccessfully determined with in-line DLS, the hydrodynamic diameter, d.nm, was additionally studied with batch DLS. Five parallel measurements resulted on the average of 15,25 nm for d.nm of milk CA VI. Graph 2 shows the five, parallel measurements as a size distribution by the sample's volume. PDI for the milk isoenzyme sample was recorded to be 0,736.



Graph 2 Size distribution by volume for five parallel measurements of milk CA VI. Records 7-11 show the five individual measurements of d.nm of the milk CA VI. Record 12 shows the average distribution based on the records 7-11.

HPLC-DLS-analysis of salivary CA VI

Like milk CA VI, also salivary CA VI was analyzed using size-exclusion chromatography and light scattering. Data in the table 3 presents SLS measurement of molecular weight (kD) for three distinct peaks 1-3 in the given retention volume (ml) and standardized molecular mass. In addition, it shows the hydrodynamic radius and hydrodynamic diameter measured by in-line DLS and batch-DLS, respectively.

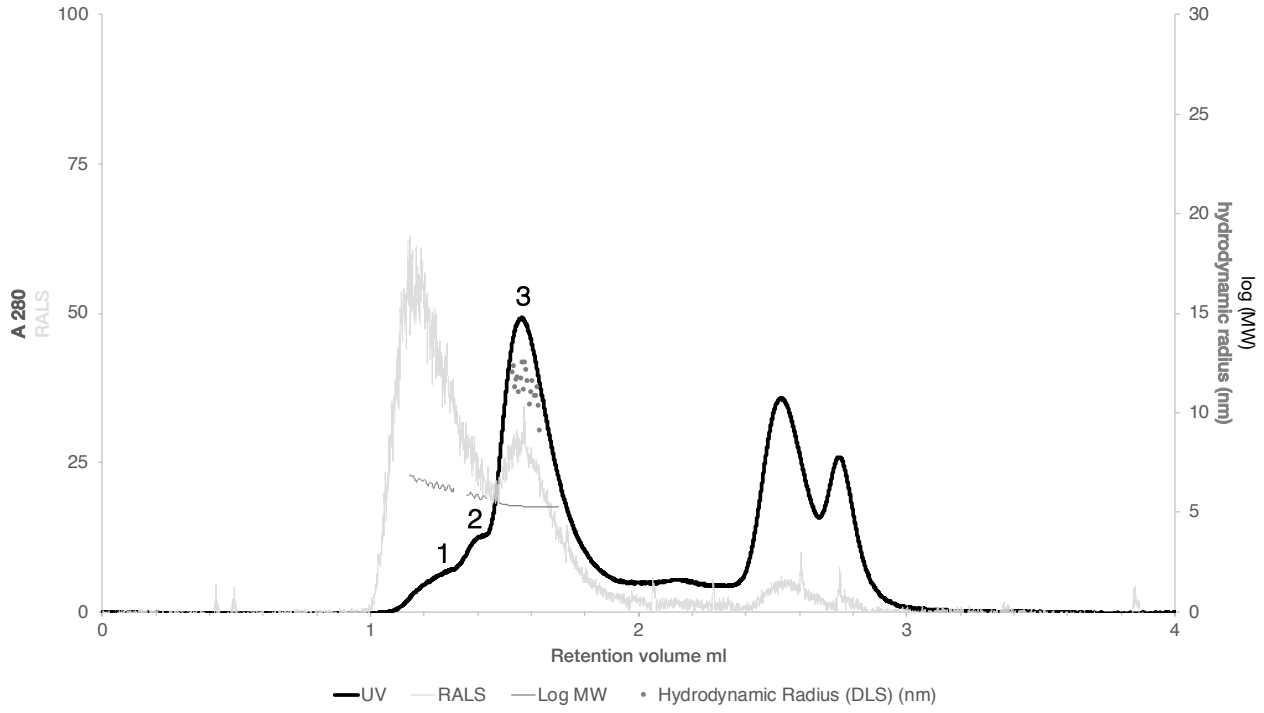
Table 3 HPLC-DLS results of salivary CA VI. In the first section, theoretical molecular weight for the monomeric human CA VI (hCA VI) protein is provided by ProtParam. In the middle section SLS and in-line DLS results, molecular weight calculated by SLS, standardized molecular weight and hydrodynamic radius, are shown for peak 1-3 in the given retention volumes. In the last section, batch DLS-measurement for hydrodynamic diameter is shown.

Sample	Theoretical MW (kD)		SEC ret. volume (ml)	SLS MW (kD)	Stand. MW (kD)	Rh DLS in-line (nm)	Batch DLS (d.nm)
salivary hCA VI	33,56952	Peak 1	1,306	2756,000	431,135	N/A	17,08
		Peak 2	1,426	675,188	292,198	N/A	
		Peak 3	**1,562**	**204,509**	**188,025**	**11,43**	

Graph 3 shows UV absorbance (A280), light scattering (RALS) and logarithmic value of molecular weight (log MW), as a function of retention volume (ml) for salivary CA VI. The

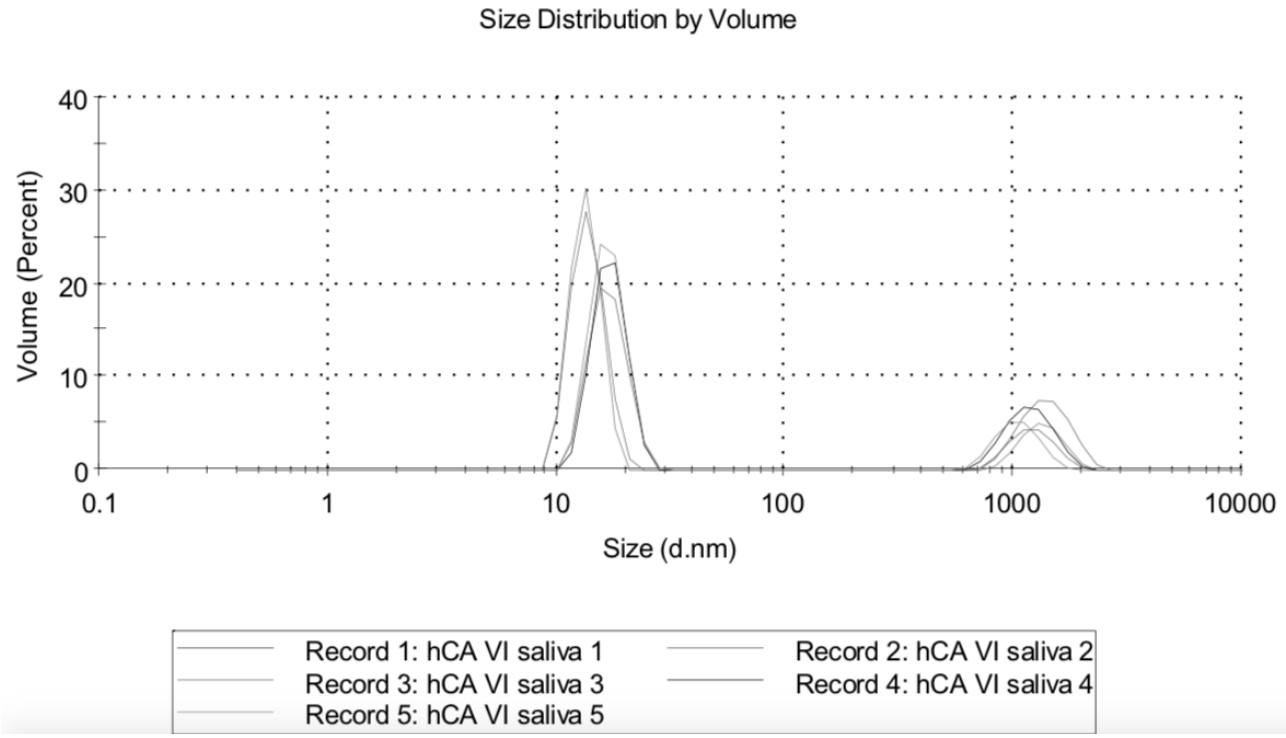
main peak, peak 3 depicted as a black curve, elutes in 1,56 ml retention volume. In retention volume 1,2 ml, a strong light scattering peak is shown as light grey as it is seen also in retention volume 1,56 ml but in lower intensity. Based on SLS measurements, molecular weight for the particles in peak 3 was computed to be 204,509 kD as it can be seen plotted as a dark grey line across the peak 3. For peak 1 and 2, the computed molecular weights were 2756 kD and 675,188 kD, respectively. Standardized molecular weight for peak 3 was 188,025 kD, whereas for peaks 1 and 2 the values were 431,135 kD and 292,198 kD, respectively.

Particle size was measured with in-line DLS and batch DLS. As seen in the graph 3, eluted peak 3 indicated the hydrodynamic radius to be 11,43 nm measured with in-line DLS.



Graph 3 Assessment of the oligomeric size of salivary human CA VI. Size-exclusion chromatography was used to study the properties of human native, salivary CA VI. In the X-axis, the retention volume for the analyzed sample is shown. The left Y-axis presents the UV absorption intensity of 280 nm wavelength (A280, black curve) and Right Angle Light Scattering intensity (RALS, light grey). In the right Y-axis, logarithmic values for molecular weight were computed by light scattering intensity. In-line DLS-measurement for hydrodynamic radius Rh is presented as a scatter plot also shown in the Y-axis.

The hydrodynamic diameter was also studied with batch DLS. Results are seen in graph 5, where the individual measurements for determining the hydrodynamic diameter are seen. Five parallel measurements resulted in average d.nm of 17,08 nm for salivary CA VI isoenzyme. PDI for the given salivary CA VI sample was recorded to be 0,481.



Graph 4 Size distribution by volume for five parallel measurements of milk CA VI. Records 1-5 show the five individual measurements of the salivary CA VI.

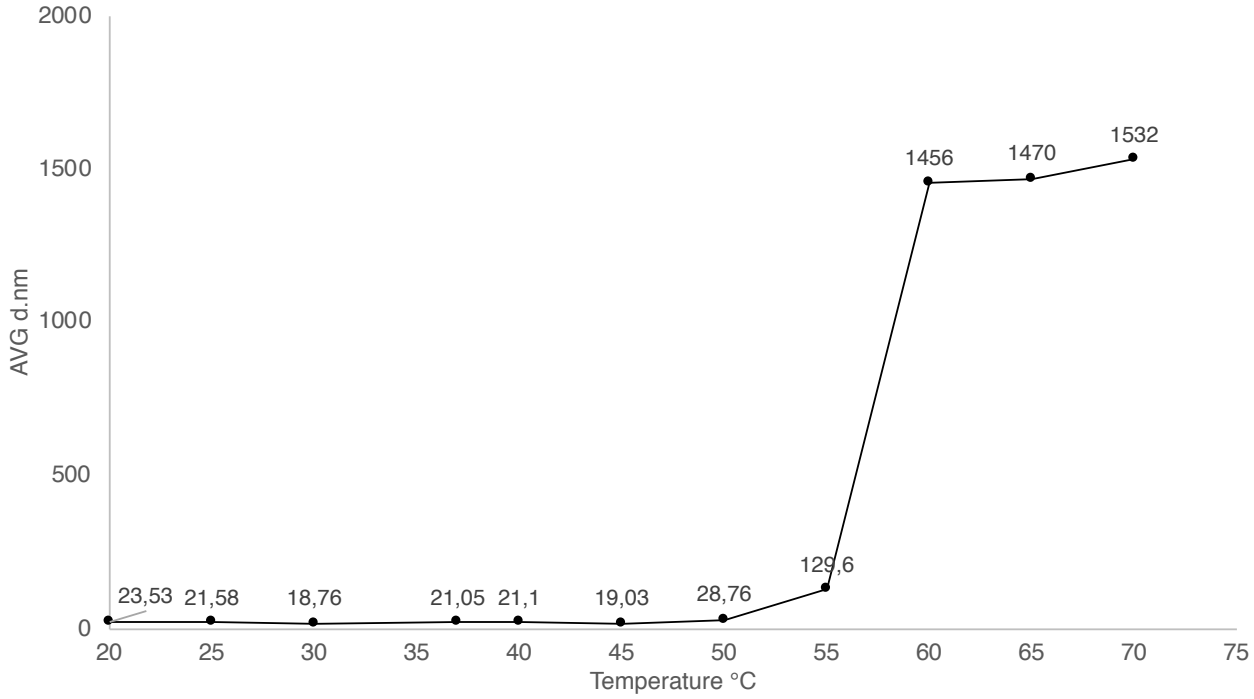
Batch DLS

The hydrodynamic diameter of salivary CA VI isoenzyme was estimated with DLS analysis. Table 4 shows the average d.nm, standard deviation for d.nm, the average percentage of the studied peak area, average PDI and standard deviation for PDI, for each of the five parallel measurements in each temperature, respectively.

Table 4 The measured values of batch-DLS-analysis. Salivary CA VI was used for determining the melting point for the studied protein. Five parallel measurements of hydrodynamic diameter were used to determine the average d.nm (average d.nm CA VI peak), the average proportional area of the studied peak (Area Vol (%) CA VI peak) and average PDI. In addition, their standard deviations were calculated, respectively (STD d.nm and STD PDI).

Temperature (°C)	Average d.nm CA VI peak (nm)	STD d.nm	Average area Vol. CA VI peak (%)	Average PDI	Standard deviation PDI
20	23,53	0,496960763	100	0,458	0,005612486
25	21,58	0,66543219	100	0,487	0,061924955
30	18,76	2,881773759	17,1	0,433	0,075400928
37	21,05	0,998839326	100	0,479	0,020916501
40	21,1	3,756776544	100	0,455	0,05157034
45	19,03	2,78890301	100	0,431	0,065078414
50	28,76	6,110864914	100	0,384	0,11361558
55	129,6	27,77314926	88,6	0,213	0,004774935
60	1456	682,5727009	0,3	0,514	0,059038123
65	1470	91,903754	100	0,279	0,062617889
70	1532	91,69896401	100	0,489	0,070604532

In the graph 5, the average value of d.nm is presented as the temperature is steadily increased from +20 °C to +70°C.



Graph 5 The hydrodynamic diameter of human salivary CA VI in given temperature points. The resulting average value for each measurement (n=5) is shown in the graph. X-axis shows the temperature in °C and Y-axis presents the hydrodynamic diameter in nm.

Throughout the data, a strong peak was detected at 19-22 nm for the hydrodynamic diameter. When reaching +50 °C, the hydrodynamic diameter changes into the micrometer scale, as seen the significant change in +50°C-+60°C range. A clear precipitation was also observed after removing the cuvette from the chamber. At the range of +25°C and +45°C, the average d.nm was 20,3 ± 1,3 nm, and the average polydispersity index 0,411 ± 0,094.

Mass spectrometry

Mass spectrometric studies were planned to perform in collaboration with Prof. Janne Jänis at the university of Eastern Finland. However, due to technical challenges, the complete mass spectrometry analyses with detailed experiments were postponed to the near future. Interestingly, the first results indicated certain differences in post-translational modifications of milk and salivary CA VI forms (Janne Jänis, personal communication, July 6th, 2018). However, more detailed analysis of the mass spectrometric data is still underway. Even though preliminary results considering the milk and salivary CA VI glycosylations were retrieved, additional information about sample preparation or the given apparatus and used methods are lacking. Thus, these results cannot be presented or discussed as it would not be considered as a good scientific practice.

Discussion

All the experiments and their results are critically discussed in the following subchapters, respectively. First, the SDS-PAGE analysis of milk and salivary CA VI are discussed in terms of molecular weight. Secondly, HPLC-DLS results are discussed considering the measured molecular weight in physiological conditions and the particle size of milk and salivary CA VI. Finally, batch DLS results, considering the size determination of salivary CA VI as well as studying the melting point of salivary CA VI, are discussed.

SDS-PAGE-analysis

After the second purification, the salivary CA VI band remained blurry, even though it was significantly purified in comparison with the first result. This tendency of salivary CA VI can be seen in the picture 1, as the salivary CAVI-band is not as well-defined and sharp as it was in milk CA VI. Milk CA VI-band possesses a sharper and stricter appearance suggesting to a less extensive deglycosylation than in salivary CA VI. Band for salivary CA VI shares this characteristic appearance with the previously reported publications of salivary CA VI in Western blot -analysis (Murakami and W S Sly, 1987; Karhumaa *et al.*,

2001; Leinonen *et al.*, 2001; Kimoto *et al.*, 2006). The major band is seen at the 42 kD, and the weaker band at 36 kD represents the deglycosylated form of salivary CA VI (Karhumaa *et al.*, 2001). The deglycosylated molecular weight was previously demonstrated as both CA VI proteins were digested with PNGase F and then, the molecular weights were shown to be identical (Karhumaa *et al.*, 2001). For milk CA VI, deglycosylations are thought to be fewer as mentioned and thus, protein is seen solely at 42 kD as presented in figure 1.

Based on the previous experiences of our research group, the affinity chromatography can yield mg quantities of carbonic anhydrase per one liter of either saliva or milk. Considering that both of the CA VI isoenzymes used in the study were native and not recombinantly produced, the amount of impurities was successfully reduced. Still, it has to be accounted that as long as some impurities exist within the studied protein, there are contaminants present that might affect the subsequent analysis.

Since the research is continued using mass spectrometry, it will be interesting to see whether the glycosylations differ. As the proteins are translated from the same gene these would be attributable to variations in post-translational modifications. If the glycosylation is reasonably consistent, it should theoretically produce a coherent and sharp band on SDS-PAGE gel. This raises the question whether the glycosylations in salivary CA VI are less stable than in milk CA VI. However, it has to be noted that oligosaccharides don't respond to the electric current like proteins and thus, glycoproteins are usually seen as more diffused in SDS-PAGE (Seppo Parkkila, personal communication, July 2nd, 2018). Since they do not assemble perfectly into spheroid form and might possess charged areas within the molecule, they appear as a blur on the gel. This makes it challenging to argue the absolute glycosylation differences between milk and salivary CA VI isoenzymes. As this research is continued, it would be suggested analyzing milk and salivary CA VI with a looser gel, e.g. 8%, so that resolution would improve, and sharper bands would be seen. Also, the mass spectrometric identification would be desirable to perform. To conclude, these analyses would offer additional information about milk and salivary CA VI post-translational structures.

HPLC-DLS

Both milk and salivary CA VI eluted in the same volume, 1,56 ml, suggesting to the similar shape and size in overall structure. In both cases, the large RALS peak (graph 1 and 3, grey curve) eluting before the main peak at 1,1 ml indicates high scattering intensity that is due to presence of aggregating protein. However, as the UV absorbance (graph 1 and 3, black curve) of the protein shows, this is assumed for only a minor percentage of the sample.

Molecular weight was determined to be nearly identical by BSA calibration; milk CA VI estimated to be 203,815 kD and salivary CA VI 204,509 kD. As their sizes were calculated by protein standard, the same retention volume resulted the same size estimation of 188,025 kD for both proteins. This estimate being slightly smaller than the SLS approximation would possibly result from the proteins' oligomeric assembly or interactions with the column matrix that would ultimately affect its penetration due to irregular shape. As the monomeric CA VI has a molecular weight of 42 kD, the HPLC measurements hint to oligomeric assembly for the CA VI. To assess the size more carefully, particle size was measured with both in-line and batch DLS. As the graphs 1 and 3 depict, the eluting peak 3 for both samples measured the salivary hydrodynamic radius to be 11,43 nm with in-line DLS but for the hydrodynamic radius of milk CA VI, reliable results were not collected (STD 39,79). For batch DLS however, hydrodynamic diameter for milk CA VI was measured to be 15,25 nm and for salivary CA VI 17,08 nm. However, the PDI for the milk CA VI was high, 0,736, implying to great polydispersity within the sample. For salivary CA VI, PDI of 0,481 implies a moderately broad polydispersity of the sample. Since high polydispersity indicates values being further apart from each other, the salivary CA VI measurement offers a more reliable result for CA VI hydrodynamic diameter as it has a lower PDI compared to milk CA VI.

As the molecular weight can be estimated by Calculator-tool in Zetasizer 7.03-software, the computed molecular weight for the hydrodynamic diameter of 17,08 nm is 508,9 kD for a globular protein. Since the additional, adsorbed structures on the surface and non-spherical form are known to affect the measured hydrodynamic diameter (Shaw, 2018), it has to be emphasized that the measured hydrodynamic diameter for CA VI, that is potentially non-spheroid and has glycosylations within its structure, might appear considerably larger in comparison with the absolute core of the protein. Thus, it has to be noted that this DLS-based value is generalized estimation made *in silico* and is not considered as relevant as the molecular weight measured with SLS as the result can be affected by the irregular shape of the oligomer. However, it has to be noted that apparently larger particle size is somewhat desired result considering that CA VI is thought to be non-spherical and has protruding surface structures.

Based on the molecular weight estimations and the large hydrodynamic radius, there is rather strong evidence to argue that human CA VI assembles into a pentameric state in a physiological buffer. To date, the pentameric structure has not been reported. The pentameric result is still plausible with the SDS-PAGE-result. Due to the extreme conditions in SDS-PAGE sample preparation, CA VI is reduced into a monomeric state because of the high temperature and the presence of mercaptoethanol in the loading buffer. Generally, only the monomeric band is seen in the SDS-PAGE-gel.

Despite the several attempts, the complete and native CA VI has not been managed to crystallize successfully. To date, only the catalytic domain of human CA VI, including amino acids 32-279, was crystallized (Pilka *et al.*, 2012). It was reported that amino acids 21-31 were disordered and, thus, excluded from the computed, structural model (Pilka *et al.*, 2012). Dimeric tendency for recombinantly produced human CA VI was reported as SEC analysis showed a two-peaked profile of CA VI suggesting to a presence of a dimeric CA VI as well as monomeric CA VI (Pilka *et al.*, 2012). The published SEC-graph by Pilka *et al.* has somewhat altered characteristics than the graphs 1 and 3 showing three peaks of which the peak 3 is the largest by its area. Pilka *et al.* show a graph with the nearly equal peaks of which the other is estimated to have the same molecular weight than 'molecular standard' 66 kD. Pilka *et al.* provided no additional data to support the dimeric nature as SEC was eventually used for protein purification in the described study. Still, Pilka *et al.* used recombinantly produced CA VI that differs from the native CA VI as it lacks amino acids from both N- and C-termini. In the given study, salivary and milk CA VI were native proteins and thus, these results cannot be compared as the studied structure is different. In conclusion, the suggested pentameric assembly that was discovered provides new information about native CA VI protein assembly. Also, the diffraction data of full-length CA VI would offer essential information considering the complete structure of CA VI and describe the oligomerization sites of CA VI as well.

The dimeric tendency of sheep lacrimal CA VI was briefly reported based on Western analysis (Ogawa *et al.*, 2002). Yet, no additional experiments to support the finding were shown.

As pentameric CA VI has not been reported previously, this novel assembly proposes new aspects considering the functionality of CA VI. Pentameric structure for CA VI has been discovered in CA VI found in zebra fish since it is found to form a complex with the pentraxin domain and subsequently form a pentamer of those complexes (Patrikainen *et al.*, 2017). As it is highly expressed in tissues forming the physical barrier against external environment, e.g. gut, skin and gills, it was thought to be a component of the innate immune system (Patrikainen *et al.*, 2017). Human CA VI differs from CA VI-PTX-complex as it seems to form a pentameric structure solely. Thus, the functional meaning of this assembly is currently unclear. Interestingly, the pentameric assembly was retained in human body temperature, +37 °C, indicating that CA VI is able to sustain its form in physiological conditions as well as in body temperature. As the research is continued, the monomeric CA VI could be studied by its structure *in silico* to evaluate potential oligomerization sites. It would potentially explain how the given pentamer is formed. Also, the pentameric assembly could be verified by performing the same analyses with recombinantly produced protein.

HPLC-DLS-analysis provided the desired data considering the molecular weight and hydrodynamic size of milk and salivary CA VI isoenzymes. For sample preparation, it would have been appropriate to follow a more detailed protocol in order to minimize measuring errors. The concentration values of protein samples were not adjusted for the specific value of CA VI extinction factor resulting in slightly greater values and thus, the absolute amount of protein injected in the HPLC-DLS-analysis was slightly less than 50 μg in both samples. This might partially explain the unsuccessful in-line DLS-measurement for milk CA VI, as there might have been an insufficient number of CA VI isoenzyme particles for determining the hydrodynamic radius. Also, the injected samples were not thoroughly centrifuged after collecting the samples from concentrating tubes but briefly spinned before transferring into HPLC-autosampler vials. Also, samples were not filtered with 0,2 μm filter. This might have caused the presence of the aggregating protein as seen in the graphs 1 and 3 as a high peak of RALS (grey curve) in ~1,1 ml retention volume. To avoid this, filtering would be advised to perform or alternatively the sample could be thoroughly centrifuged and transferred into a new tube. Due to the complex shape of the UV-absorbance curve in graphs 1 and 3, all the eluted proteins of different peaks could additionally be fractioned and subsequently analyzed with SDS-PAGE to reassure their molecular weight.

For good scientific practice, it would be appropriate to repeat the complete analysis with a more detailed sample preparation described previously to verify the measured values as the given experiment was performed only once for each sample. In addition, it would be intriguing to analyze both salivary and milk CA VI in SEC with different temperatures to observe how the temperature would affect its oligomeric assembly.

Batch DLS

Salivary CA VI was more carefully studied as the result for hydrodynamic diameter was verified in a separate measurement. As seen in the graph 3, average d.nm for salivary CA VI between 25-45 °C is 20,304 ± 1,3 nm. Scattering intensity and d.nm remain constant as shown in the table 4. Similarly, PDI was considered stable between 25-45 °C, average PDI being 0,457±1,325. However, the percentage of the studied peak area was 100% in all the measured temperatures except +30°C, +55 °C and +60 °C (17,1%, 88,6% and 0,3% respectively) showing that majority of the measurements recorded the hydrodynamic radius of CA VI accurately since the resulting diameter was highly similar in comparison with the hydrodynamic radius measured with SLS. Since there are no previous reported values for the hydrodynamic size of CA VI, this result cannot be compared to the preexisting data.

The measured hydrodynamic diameter for salivary CA VI was drastically increased at +55 °C indicating the presence of denatured CA VI aggregates, hence, the melting point for salivary CA VI must reside between +50 - +55 °C. This result supports the formerly measured value, +53,6 °C (Kazokaite *et al.*, 2015), for the melting temperature of native, salivary CA VI isoenzyme.

Since both of the studied CA VI isoenzymes are native proteins, the impurities will affect adversely by increasing the polydispersity of the analyzed samples. As the percentage of the peak area in +30°C was only 17%, it would be advisable to re-determine the given value for large contaminants. Thus, to lower the polydispersity index of the samples, the amount of impurities should be reduced even further. Also, significantly differing values could be excluded from the data for lowering the PDI.

Due to insufficient amount of milk CA VI, the batch DLS for the given isoenzyme was not performed in this study. For the upcoming experiments, similar batch DLS analysis for milk CA VI would provide reassuring data for the melting point of salivary CA VI. It would also confirm that both milk and salivary CA VI would possess the same melting temperature, or, whether the glycosylation differences would have an effect on the given parameter. For recombinantly produced CA VI, both in *E. Coli* and in mammalian cells, the melting temperature was shown to be slightly lower than for native, salivary CA VI (Kazokaite *et al.*, 2015). This was assumed to depend on glycosylation differences (Kazokaite *et al.*, 2015). Hence, it would be recommended to analyze milk and salivary CA VI isoenzymes in parallel manner as this study is continued in the near future.

Conclusion

To conclude this study, milk and salivary CA VI were purified with affinity chromatography, and SDS-PAGE analysis was parallelly performed to assess their molecular weight and glycosylation differences. In SDS-PAGE analysis salivary CA VI was seen as more dispersed than milk CA VI, suggesting less extensive deglycosylation in milk CA VI. This variation between the milk and salivary CA VI could be more properly studied with MS identification.

The molecular weight and particle size were estimated with HPLC-DLS and batch DLS, respectively. SLS determined the milk CA VI to have a molecular weight of 203,815 kD and 204,509 kD for salivary CA VI. Standardized molecular weight for both milk and salivary CA VI was calculated to be 188,025 kD. The hydrodynamic radius for milk CA VI was not retrieved from in-line DLS, but from batch DLS, the hydrodynamic diameter was measured to be 15,25 nm. For salivary CA VI, the hydrodynamic radius from in-line DLS was measured

to be 11,43 nm and batch DLS measured the hydrodynamic diameter to be 17,08 nm. The hydrodynamic diameter of salivary CA VI was additionally determined with increasing temperature and found to have an average of 20,3 ± 1,3 nm between temperatures +25°C and +45°C. The great particle size and molecular weight indicate that CA VI assembles as an oligomer, a pentameric structure being the most probable form. To date, this pentameric structure has not been reported. This novel form offers new insights for CA VI functions that need to be further studied. Lastly, the range for melting point of the salivary CA VI was measured to be +50 - +55 °C. The result was consistent with the previously reported value. MS analyses are underway to determine potential differences in post-translational modifications, such as glycosylation, in both milk and salivary CA VI enzymes.

References

Aidar, M. *et al.* (2013) 'Effect of genetic polymorphisms in CA6 gene on the expression and catalytic activity of human salivary carbonic anhydrase VI', *Caries research*, 47(5), pp. 414–420. doi: 10.1159/000350414 [doi].

Alterio, V. *et al.* (2009) 'Crystal structure of the catalytic domain of the tumor-associated human carbonic anhydrase IX.', *Proceedings of the National Academy of Sciences of the United States of America*, 106(38), pp. 16233–16238. doi: 10.1073/pnas.0908301106.

Asari, M. *et al.* (2000) 'Distribution of carbonic anhydrase isozyme VI in the developing bovine parotid gland', *Cells, Tissues, Organs*, 167(1), pp. 18–24. doi: 10.1159/000016762.

Aspatwar, A., Tolvanen, M. E. and Parkkila, S. (2010) 'Phylogeny and expression of carbonic anhydrase-related proteins', *BMC molecular biology*, 11, p. 25. doi: 10.1186/1471-2199-11-25 [doi].

Barbarossa, I. T. *et al.* (2015) 'The gustin (CA6) gene polymorphism, rs2274333 (A/G), is associated with fungiform papilla density, whereas PROP bitterness is mostly due to TAS2R38 in an ethnically-mixed population', *Physiology and Behavior*, 138, pp. 6–12. doi: 10.1016/j.physbeh.2014.09.011.

Barker, H. *et al.* (2017) 'Role of carbonic anhydrases in skin wound healing', *Experimental & molecular medicine*, 49(5), p. . Epub 2017 May 19 doi:10.1038/emm.2017.60. doi: 10.1038/emm.2017.60 [doi].

Barnett, D. H. *et al.* (2008) 'Estrogen receptor regulation of carbonic anhydrase XII through a distal enhancer in breast cancer', *Cancer Research*, 68(9), pp. 3505–3515. doi: 10.1158/0008-5472.CAN-07-6151.

Bootorabi, F. *et al.* (2010) 'Analysis of a shortened form of human carbonic anhydrase VII expressed in vitro compared to the full-length enzyme', *Biochimie*, 92(8), pp. 1072–1080. doi: 10.1016/j.biochi.2010.05.008.

Bottazzi, B. *et al.* (2016) 'The pentraxins PTX3 and SAP in innate immunity, regulation of inflammation and tissue remodelling', *Journal of Hepatology*, 64(6), pp. 1416–1427. doi: 10.1016/j.jhep.2016.02.029.

Brown, D. *et al.* (1983) 'Immunohistochemical localization of carbonic anhydrase in postnatal and adult rat kidney', *The American Journal of Physiology*, 245(1), p. 110. doi:

10.1152/ajprenal.1983.245.1.F110.

Calò, C. *et al.* (2011) 'Polymorphisms in TAS2R38 and the taste bud trophic factor, gustin gene co-operate in modulating PROP taste phenotype', *Physiology and Behavior*, 104(5), pp. 1065–1071. doi: 10.1016/j.physbeh.2011.06.013.

Chegwidden, W. R. and Carter, N. D. (2000) 'Introduction to the carbonic anhydrases', *EXS*, (90)(90), pp. 14–28.

Culp, D. J. *et al.* (2011) 'Oral colonization by Streptococcus mutans and caries development is reduced upon deletion of carbonic anhydrase VI expression in saliva', *Biochimica et biophysica acta*, 1812(12), pp. 1567–1576. doi: 10.1016/j.bbadis.2011.09.006 [doi].

Feeney, E. L. and Hayes, J. E. (2014) 'Exploring associations between taste perception, oral anatomy and polymorphisms in the carbonic anhydrase (gustin) gene CA6', *Physiology & Behavior*, 128, pp. 148–154. doi: 10.1016/j.physbeh.2014.02.013.

Feldstein, J. B. and Silverman, D. N. (1984) 'Purification and characterization of carbonic anhydrase from the saliva of the rat', *The Journal of Biological Chemistry*, 259(9), pp. 5447–5453.

Fernley, R. T., Coghlan, J. P. and Wright, R. D. (1988) 'Purification and characterization of a high-Mr carbonic anhydrase from sheep parotid gland', *The Biochemical Journal*, 249(1), pp. 201–207.

Fernley, R. T., Wright, R. D. and Coghlan, J. P. (1979) 'A novel carbonic anhydrase from the ovine parotid gland', *FEBS letters*, 105(2), pp. 299–302.

Frost, S. C. (2014) 'Physiological Functions of the Alpha Class of Carbonic Anhydrases', *Carbonic Anhydrase: Mechanism, Regulation, Links to Disease, and Industrial Applications*, pp. 9–30. doi: 10.1007/978-94-007-7359-2_2.

Gasteiger, E. *et al.* (2005) 'Protein Identification and Analysis Tools on the ExPASy Server', *The Proteomics Protocols Handbook*. doi: 10.1385/1-59259-890-0:571.

Guinez, C. *et al.* (2008) 'Protein ubiquitination is modulated by O-GlcNAc glycosylation', *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 22(8), pp. 2901–2911. doi: 10.1096/fj.07-102509.

Henkin, R. I., Martin, B. M. and Agarwal, R. P. (1999a) 'Decreased parotid saliva

gustin/carbonic anhydrase VI secretion: an enzyme disorder manifested by gustatory and olfactory dysfunction', *The American Journal of the Medical Sciences*, 318(6), pp. 380–391.

Henkin, R. I., Martin, B. M. and Agarwal, R. P. (1999b) 'Efficacy of Exogenous Oral Zinc in Treatment of Patients with Carbonic Anhydrase VI Deficiency', *The American Journal of the Medical Sciences*, 318(6), pp. 392–405. doi: 10.1016/S0002-9629(15)40664-0.

Hooper, L. V, Hindsgaul, O. and Baenziger, J. U. (1995) 'Purification and characterization of the GalNAc-4-sulfotransferase responsible for sulfation of GalNAc beta 1,4GlcNAc-bearing oligosaccharides', *J Biol Chem*, 270(27), pp. 16327–16332. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7608201.

Innocenti, A., Scozzafava, A. and Supuran, C. T. (2010) 'Carbonic anhydrase inhibitors. Inhibition of transmembrane isoforms IX, XII, and XIV with less investigated anions including trithiocarbonate and dithiocarbamate', *Bioorganic & medicinal chemistry letters*, 20(5), pp. 1548–1550. doi: 10.1016/j.bmcl.2010.01.081 [doi].

Jonasz, M. and Fournier, G. R. (2007) *Light Scattering by Particles in Water: Theoretical and Experimental Foundations*, *Light Scattering by Particles in Water: Theoretical and Experimental Foundations*. doi: 10.1016/B978-0-12-388751-1.X5000-5.

Kalinin, S., Supuran, C. T. and Krasavin, M. (2016) 'Multicomponent chemistry in the synthesis of carbonic anhydrase inhibitors', *Journal of enzyme inhibition and medicinal chemistry*, 31(sup4), pp. 185–199. doi: 10.1080/14756366.2016.1220944 [doi].

Karhumaa, P. *et al.* (2001) 'The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk', *Proceedings of the National Academy of Sciences of the United States of America*, 98(20), pp. 11604–11608. doi: 10.1073/pnas.121172598 [doi].

Kaseda, M. *et al.* (2006) 'Immunohistochemistry of the bovine secretory carbonic anhydrase isozyme (CA-VI) in bovine alimentary canal and major salivary glands', *The Journal of Veterinary Medical Science*, 68(2), pp. 131–135.

Kasuya, T. *et al.* (2007) 'Immunohistolocalization and gene expression of the secretory carbonic anhydrase isozymes (CA-VI) in canine oral mucosa, salivary glands and oesophagus', *Anatomia, Histologia, Embryologia*, 36(1), pp. 53–57. doi: 10.1111/j.1439-0264.2006.00721.x.

Kazokaite, J. *et al.* (2015) 'Intrinsic binding of 4-substituted-2,3,5,6-tetrafluorobenezenesulfonamides to native and recombinant human carbonic anhydrase

VI', *FEBS Journal*. doi: 10.1111/febs.13196.

Kimoto, M. *et al.* (2004) 'Carbonic anhydrase VI in the mouse nasal gland', *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society*, 52(8), pp. 1057–1062. doi: 10.1369/jhc.3A6243.2004.

Kimoto, M. *et al.* (2006) 'A role of salivary carbonic anhydrase VI in dental plaque', *Archives of Oral Biology*, 51(2), pp. 117–122. doi: 10.1016/j.archoralbio.2005.04.007.

Kinnamon, S. C. and Roper, S. D. (1987) 'Passive and active membrane properties of mudpuppy taste receptor cells', *Journal of Physiology*, 383, pp. 601–614.

Kitade, K. *et al.* (2003) 'Expression and localization of carbonic anhydrase in bovine mammary gland and secretion in milk', *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, 134(2), pp. 349–354.

Kivelä, J., Parkkila, S., Metteri, J., *et al.* (1997) 'Salivary carbonic anhydrase VI concentration and its relation to basic characteristics of saliva in young men', *Acta Physiologica Scandinavica*, 161(2), pp. 221–225. doi: 10.1046/j.1365-201X.1997.00217.x.

Kivelä, J., Parkkila, S., Waheed, A., *et al.* (1997) 'Secretory carbonic anhydrase isoenzyme (CA VI) in human serum', *Clinical Chemistry*, 43(12), pp. 2318–2322.

Kivelä, J., Parkkila, S., Parkkila, A. K. and Rajaniemi, H. (1999) 'A low concentration of carbonic anhydrase isoenzyme VI in whole saliva is associated with caries prevalence', *Caries Research*, 33(3), pp. 178–184. doi: 10.1159/000016514.

Kivelä, J., Parkkila, S., Parkkila, A. K., Leinonen, J., *et al.* (1999) 'Salivary carbonic anhydrase isoenzyme VI', *The Journal of Physiology*, 520 Pt 2, pp. 315–320.

Kivelä, J. *et al.* (2003) 'Salivary carbonic anhydrase VI and its relation to salivary flow rate and buffer capacity in pregnant and non-pregnant women', *Archives of Oral Biology*, 48(8), pp. 547–551. doi: 10.1016/S0003-9969(03)00096-7.

Law, J. S. *et al.* (1987) 'Human salivary gustin is a potent activator of calmodulin-dependent brain phosphodiesterase.', *Proceedings of the National Academy of Sciences of the United States of America*, 84(6), pp. 1674–8. doi: 10.1073/pnas.84.6.1674.

Lehtonen, J. *et al.* (2004) 'Characterization of CA XIII, a novel member of the carbonic anhydrase isozyme family', *The Journal of Biological Chemistry*, 279(4), pp. 2719–2727. doi: 10.1074/jbc.M308984200 [doi].

Leinonen, J. *et al.* (1999) 'Salivary Carbonic Anhydrase Isoenzyme VI Is Located in the Human Enamel Pellicle', *Caries Research*, 33(3), pp. 185–190. doi: 10.1159/000016515.

Leinonen, J. *et al.* (2001) 'Secretion of carbonic anhydrase isoenzyme VI (CA VI) from human and rat lingual serous von Ebner's glands', *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society*, 49(5), pp. 657–662. doi: 10.1177/002215540104900513.

Leinonen, J. (2008) *Carbonic anhydrase isoenzyme VI: distribution, catalytic properties and biological significance*. University of Oulu. Available at: http://jultika.oulu.fi/files/isbn9789514289903.pdf.

Leinonen, J. S. *et al.* (2004) 'Immunohistochemical demonstration of carbonic anhydrase isoenzyme VI (CA VI) expression in rat lower airways and lung', *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society*, 52(8), pp. 1107–1112. doi: 10.1369/jhc.4A6282.2004.

Lönnerholm, G. and Wistrand, P. J. (1983) 'Carbonic anhydrase in the human fetal kidney', *Pediatric Research*, 17(5), pp. 390–397. doi: 10.1203/00006450-198305000-00015.

Malvern Instruments Ltd. (2008) *Zetasizer Nano Series user manual*. Worcestershire.

Malvern Instruments Ltd. (2014) 'Static Light Scattering technologies for GPC-SEC explained', *Malvern Instruments*. doi: papers3://publication/uuid/9675F9A0-F391-4B89-B7D5-E59B979703EC.

Matossian, C. and Micucci, J. (2016) 'Characterization of the serological biomarkers associated with sjögren's syndrome in patients with recalcitrant dry eye disease', *Clinical Ophthalmology*, pp. 1329–1334. doi: 10.2147/OPTH.S106973.

Matthews, T. A. *et al.* (2014) 'Expression of the CHOP-inducible carbonic anhydrase CAVI-b is required for BDNF-mediated protection from hypoxia', *Brain Research*. doi: 10.1016/j.brainres.2013.11.018.

McCormick, N. H. and Kelleher, S. L. (2012) 'ZnT4 provides zinc to zinc-dependent proteins in the trans-Golgi network critical for cell function and Zn export in mammary epithelial cells', *American Journal of Physiology. Cell Physiology*, 303(3), p. 291. doi: 10.1152/ajpcell.00443.2011.

Melis, M. *et al.* (2013) 'The gustin (CA6) gene polymorphism, rs2274333 (A/G), as a mechanistic link between PROP tasting and fungiform taste papilla density and

maintenance', *PloS One*, 8(9), p. e74151. doi: 10.1371/journal.pone.0074151.

Moldoveanu, S. and David, V. (2013) *Essentials in Modern HPLC Separations*, *Essentials in Modern HPLC Separations*. doi: 10.1016/C2010-0-65748-8.

Montgomery, J. C. *et al.* (1991) 'Characterization of the human gene for a newly discovered carbonic anhydrase, CA VII, and its localization to chromosome 16', *Genomics*, 11(4), pp. 835–848.

Murakami, H. and Sly, W. S. (1987) 'Purification and characterization of human salivary carbonic anhydrase', *Journal of Biological Chemistry*, 262(3), pp. 1382–1388.

Murakami, H. and Sly, W. S. (1987) 'Purification and characterization of human salivary carbonic anhydrase', *The Journal of Biological Chemistry*, 262(3), pp. 1382–1388.

Netopilík, M. (2017) 'Toward ideal separation by size-exclusion chromatography', *Journal of Chromatography A*. doi: 10.1016/j.chroma.2017.01.038.

Nishimori, I. *et al.* (2007) 'Carbonic anhydrase inhibitors. Inhibition studies of the human secretory isoform VI with anions', *Bioorganic and Medicinal Chemistry Letters*, 17(4), pp. 1037–1042. doi: 10.1016/j.bmcl.2006.11.028.

Nishita, T. *et al.* (2001) 'Purification of carbonic anhydrase isozyme VI (CA-VI) from swine saliva', *The Journal of Veterinary Medical Science*, 63(10), pp. 1147–1149.

Nishita, T. *et al.* (2007) 'Measurement of carbonic anhydrase isozyme VI (CA-VI) in bovine sera, saliva, milk and tissues', *Veterinary Research Communications*, 31(1), pp. 83–92. doi: 10.1007/s11259-006-3423-0.

Nishita, T. *et al.* (2011) 'Measurement of carbonic anhydrase isozyme VI (CA-VI) in swine sera, colostrums, saliva, bile, seminal plasma and tissues', *Animal Science Journal = Nihon Chikusan Gakkaiho*, 82(5), pp. 673–678. doi: 10.1111/j.1740-0929.2011.00888.x.

Nishita, T., Yatsu, J., Murakami, M., *et al.* (2014) 'Isolation and sequencing of swine carbonic anhydrase VI, an enzyme expressed in the swine kidney', *BMC research notes*, 7, p. 116. doi: 10.1186/1756-0500-7-116.

Nishita, T., Yatsu, J., Watanabe, K., *et al.* (2014) 'Urinary carbonic anhydrase VI as a biomarker for kidney disease in pigs', *Veterinary Journal (London, England: 1997)*, 202(2), pp. 378–380. doi: 10.1016/j.tvjl.2014.07.007.

Nobbmann, U. (2014) *Polydispersity – what does it mean for DLS and chromatography?* Available at: http://www.materials-talks.com/blog/2014/10/23/polydispersity-what-does-it-mean-for-dls-and-chromatography/.

Ogawa, Y. *et al.* (2002) 'Characterization of lacrimal gland carbonic anhydrase VI', *Journal of Histochemistry and Cytochemistry*, 50(6), pp. 821–827. doi: 10.1177/002215540205000608.

Padiglia, A. *et al.* (2010) 'Sensitivity to 6-n-propylthiouracil is associated with gustin (carbonic anhydrase VI) gene polymorphism, salivary zinc, and body mass index in humans', *The American Journal of Clinical Nutrition*, 92(3), pp. 539–545. doi: 10.3945/ajcn.2010.29418.

Pan, P. W. *et al.* (2011) 'Gene expression profiling in the submandibular gland, stomach, and duodenum of CAVI-deficient mice', *Transgenic research*, 20(3), pp. 675–698. doi: 10.1007/s11248-010-9441-2 [doi].

Panchal, J. *et al.* (2014) 'Analyzing Subvisible Particles in Protein Drug Products: a Comparison of Dynamic Light Scattering (DLS) and Resonant Mass Measurement (RMM)', *The AAPS Journal*. doi: 10.1208/s12248-014-9579-6.

Parkkila, S. *et al.* (1990) 'Immunohistochemical localization of carbonic anhydrase isoenzymes VI, II, and I in human parotid and submandibular glands', *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society*, 38(7), pp. 941–947. doi: 10.1177/38.7.2113069.

Parkkila, S., Parkkila, A. K. and Rajaniemi, H. (1995) 'Circadian periodicity in salivary carbonic anhydrase VI concentration', *Acta Physiologica Scandinavica*, 154(2), pp. 205–211. doi: 10.1111/j.1748-1716.1995.tb09902.x.

Patrikainen, M. *et al.* (2014) 'The role of carbonic anhydrase VI in bitter taste perception: evidence from the Car6(-)/(-) mouse model', *Journal of Biomedical Science*, 21, p. 2. doi: 10.1186/s12929-014-0082-2 [doi].

Patrikainen, M. S. *et al.* (2016) 'Altered gene expression in the lower respiratory tract of Car6 (-/-) mice', *Transgenic research*, 25(5), pp. 649–664. doi: 10.1007/s11248-016-9961-5 [doi].

Patrikainen, M. S. *et al.* (2017) 'Identification and characterization of a novel zebrafish (Danio rerio) pentraxin-carbonic anhydrase', *PeerJ*, 5, p. e4128. doi: 10.7717/peerj.4128.

Peres, R. C. R. *et al.* (2010) 'Association of polymorphisms in the carbonic anhydrase 6 gene with salivary buffer capacity, dental plaque pH, and caries index in children aged 7–9 years', *The Pharmacogenomics Journal*, 10(2), p. 114. doi: 10.1038/tpj.2009.37.

Pertovaara, M. *et al.* (2011) 'Novel carbonic anhydrase autoantibodies and renal manifestations in patients with primary Sjogren's syndrome.', *Rheumatology (Oxford, England)*, 50(8), pp. 1453–7. doi: 10.1093/rheumatology/ker118.

Pilka, E. S. *et al.* (2012) 'Crystal structure of the secretory isozyme of mammalian carbonic anhydrases CA VI: implications for biological assembly and inhibitor development', *Biochemical and Biophysical Research Communications*, 419(3), pp. 485–489. doi: 10.1016/j.bbrc.2012.02.038.

Poulsen, S.-A. *et al.* (2008) 'Inhibition of human mitochondrial carbonic anhydrases VA and VB with para-(4-phenyltriazole-1-yl)-benzenesulfonamide derivatives', *Bioorganic & Medicinal Chemistry Letters*, 18(16), pp. 4624–4627. doi: 10.1016/j.bmcl.2008.07.010.

Rapp, G. W. (1946) 'The Biochemistry of Oral Calculus. II. The Presence of Carbonic Anhydrase in Human Saliva', *The Journal of the American Dental Association*. Elsevier, 33(3), pp. 191–194. doi: 10.14219/JADA.ARCHIVE.1946.0047.

Shah, G. N. *et al.* (2000) 'Mitochondrial carbonic anhydrase CA VB: differences in tissue distribution and pattern of evolution from those of CA VA suggest distinct physiological roles', *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp. 1677–1682.

Shaw, R. (2018) *Dynamic Light Scattering Training –Achieving reliable nano particle sizing*. Available at: http://149.171.168.221/partcat/wp-content/uploads/Malvern-Zetasizer-LS.pdf (Accessed: 29 June 2018).

Shimadzu corporation (2018) *No Title*. Available at: https://www.shimadzu.com/an/hplc/inert_20ai.html.

Slotboom, D. J. *et al.* (2008) 'Static light scattering to characterize membrane proteins in detergent solution', *Methods*. doi: 10.1016/j.ymeth.2008.06.012.

Sok, J. *et al.* (1999) 'CHOP-Dependent stress-inducible expression of a novel form of carbonic anhydrase VI.', *Molecular and cellular biology*. doi: 10.1128/MCB.19.1.495.

Stams, T. and Christianson, D. W. (2000) 'X-ray crystallographic studies of mammalian carbonic anhydrase isozymes', *EXS*, (90)(90), pp. 159–174.

Striegel, A. M. *et al.* (2009) *Modern Size-Exclusion Liquid Chromatography: Practice of Gel Permeation and Gel Filtration Chromatography: Second Edition*, *Modern Size-Exclusion Liquid Chromatography: Practice of Gel Permeation and Gel Filtration Chromatography: Second Edition*. doi: 10.1002/9780470442876.

Sumi, K. R. *et al.* (2018) 'Molecular cloning and characterization of secretory carbonic anhydrase VI in pufferfish (Takifugu rubripes)', *Gene*, 640, pp. 57–65. doi: 10.1016/j.gene.2017.10.008.

Supuran, C. T. and De Simone, G. (2015) *Carbonic Anhydrases as Biocatalysts: From Theory to Medical and Industrial Applications*, *Carbonic Anhydrases as Biocatalysts: From Theory to Medical and Industrial Applications*. doi: 10.1016/C2012-0-13548-1.

Sutherland, G. R. *et al.* (1989) 'The gene for human carbonic anhydrase VI(CA6) is on the tip of the short arm of chromosome 1.', *Cytogenetics and cell genetics*. Switzerland, 50(2–3), pp. 149–150. doi: 10.1159/000132746.

Svastova, E. *et al.* (2012) 'Carbonic anhydrase IX interacts with bicarbonate transporters in lamellipodia and increases cell migration via its catalytic domain', *The Journal of biological chemistry*, 287(5), pp. 3392–3402. doi: 10.1074/jbc.M111.286062 [doi].

Szabó, I. (1974) 'Carbonic anhydrase activity in the saliva of children and its relation to caries activity', *Caries Research*, 8(2), pp. 187–191. doi: 10.1159/000260107.

Tashian, R. E. *et al.* (2000) 'Carbonic anhydrase (CA)-related proteins (CA-RPs), and transmembrane proteins with CA or CA-RP domains', *EXS*, (90)(90), pp. 105–120.

Thatcher, B. J. *et al.* (1998) 'Gustin from human parotid saliva is carbonic anhydrase VI', *Biochemical and Biophysical Research Communications*, 250(3), pp. 635–641. doi: 10.1006/bbrc.1998.9356.

Thomas, J. A. and Mallis, R. J. (2001) 'Aging and oxidation of reactive protein sulfhydryls', *Experimental Gerontology*, 36(9), pp. 1519–1526.

Tolvanen, M. E. E. *et al.* (2013) 'Analysis of evolution of carbonic anhydrases IV and XV reveals a rich history of gene duplications and a new group of isozymes', *Bioorganic & medicinal chemistry*, 21(6), pp. 1503–1510. doi: 10.1016/j.bmc.2012.08.060 [doi].

Vullo, D. *et al.* (2017) 'Comparison of the Sulfonamide Inhibition Profiles of the beta- and gamma-Carbonic Anhydrases from the Pathogenic Bacterium Burkholderia pseudomallei', *Molecules (Basel, Switzerland)*, 22(3), p. 10.3390/molecules22030421. doi: E421 [pii].

Wang, C.-C. *et al.* (2007) 'VAMP8/endobrevin as a general vesicular SNARE for regulated exocytosis of the exocrine system', *Molecular Biology of the Cell*, 18(3), pp. 1056–1063. doi: 10.1091/mbc.E06-10-0974.

Whittington, D. A. *et al.* (2004) 'Expression, assay, and structure of the extracellular domain of murine carbonic anhydrase XIV: Implications for selective inhibition of membrane-associated isozymes', *Journal of Biological Chemistry*, 279(8), pp. 7223–7228. doi: 10.1074/jbc.M310809200.

Wolpert, H. R., Strader, C. D. and Khalifah, R. G. (1977) 'Interaction of the Unique Competitive Inhibitor Imidazole with Human Carbonic Anhydrase B', *Biochemistry*, 16(26), pp. 5717–5721. doi: 10.1021/bi00645a011.

Yang, Z. *et al.* (2005) 'Mutant carbonic anhydrase 4 impairs pH regulation and causes retinal photoreceptor degeneration', *Human Molecular Genetics*, 14(2), pp. 255–265. doi: 10.1093/hmg/ddi023.

Appendices

Appendix 1 Used solutions in protein purification

Solution	Reagents	pH
Collecting buffer for CA VI	0,1 M TRIS (VWR 0497, #040GC459) 0,2 M Na_2SO_4 (Honeywell, #BCBS8593V) 0,2 M Benzamidine (Sigma B6506, #117K0123)	8,70
Binding buffer for CA VI	0,1 M TRIS (VWR 0497, #040GC459) 0,2 M Na_2SO_4 (Honeywell, #BCBS8593V)	8,70
Wash 1 buffer for CA VI	0,1 M TRIS (VWR 0497, #040GC459) 0,2 M Na_2SO_4 (Honeywell, #BCBS8593V) 0,2 M Benzamidine (Sigma B6506, #117K0123) 20% glycerol (Sigma, #SHBH4364V)	8,70
Wash 2 buffer for CA VI	0,1 M TRIS (VWR 0497, #040GC459) 0,2 M Na_2SO_4 (Honeywell, #BCBS8593V) 0,2 M Benzamidine (Sigma B6506, #117K0123) 20% glycerol (Sigma, #SHBH4364V)	7,00
Elution buffer for CA VI	0,1 M TRIS (VWR 0497, #040GC459) 0,4 M NaN_3 (VALMISTAJA) 1mM Benzamidine (Sigma B6506, #117K0123) 20% glycerol (Sigma, #SHBH4364V)	7,0
HPLC buffer	150 mM NaCl (VWR, #17L184138) 50 mM TRIS (VWR 0497, #040GC459)	7,5
MS buffer	100 mM ammonium acetate (Fluka, #BCBJ9119V)	NA